Characterisation of *COAS2*, a candidate proto-oncogene amplified and over-expressed in human mesenchymal tumours

Erik Braarød Paulsen

Thesis for the Cand. scient. Degree in Biochemistry

2004

Department of Tumor Biology
Institute for Cancer Research
The Norwegian Radium Hospital

Department of Molecular Biosciences
Faculty of Mathematics and Natural Sciences
University of Oslo

Acknowledgments

As my supervisor, Professor Dr. Philos Ola Myklebost has introduced me to the field of cancer research. During my period as a student in his group, I have been able to learn and use a vast array of molecular biology techniques. I have had excellent working facilities, and also attended courses which otherwise would be reserved for PhD-students. I will always be grateful for his help and feedback. Thank you.

I would also like to express my sincere gratitude to Dr. Philos Leonardo Andrés Méza-Zepeda.
Having taught me most of the practical methods and background theory, he has been my day-to-day supervisor. As a full time working scientist, it is astounding how he always found time to answer all my questions without restraints. I thank him for always being there, for conclusive answers and his good humour and spirit.

Ola Myklebost's group: Jørn Henriksen, Jeanne-Marie Berner, Stine Henrichson Kresse, Magne Skårn, Lise Ramberg, Heidi Maria Namløs, Diako Hamedi, Christoph Müller and Paul Noordhuis: Thank you. Your generosity and positivity has overwhelmed me at times, and you made my life a lot easier through the tough periods when "all things failed". Jeanne-Marie made the early cDNA microarray experiments come true, and she, as Jørn, has always found time for me. Thank you: Magne, Stine and Lise for friendship, support and excellent collaboration.

The microarray core facility at The Norwegian Radium Hospital (Bjørn E. Kristiansen, Vibeke Andreassen and Signe Indahl) assisted me at all times during the microarray experiments. I would like to thank them all for their kindness and hospitality.

Senior Scientist Dr. David Warren deserves special thanks for all the help and especially for all the work he has put down to make production of COAS2 in insect cells happen. As Dr Méza Zepeda, Dr. Warren has an ability to answer politely at all times. I remark that without their aid, much of the molecular cloning would have taken me additional months to perform.

I must mention Senior Engineer Meng Yu Wang at the Department of Tumour Biology, and thank him for introducing the soft agar assay to a young student.

I would like to express my gratitude to Professor Dr. Philos Harald Stenmark and PhD Jørgen Wesche at the Department of Biochemistry at DNR, who helped me with the confocal microscopy.

Finally, all of the staff at the Department of Tumour Biology at DNR deserves a million thanks for welcoming me in an extraordinary manner. Thank you to Jeen Dhoski, the young Biomolecular Engineering apprentice who has assisted me with occasional micropreps and for her positive attitude.

There is one person to whom I would like to express my absolute respect and love; my fiancée Lise deserves the full honour of this work.

ERIK BRAARØD PAULSEN

Oslo, May 2004

Contents

Abstract

Chromosome One Amplified Sequence 2 (COAS2) is a member of the cyclophilin family of proteins (Meza Zepeda et al., 2002). It has 84% amino acid homology to *H. sapiens* cyclophilin A, also known as Peptidylprolyl Isomerase A or PPIA. *COAS2* is localised in 1q21, and frequently amplified and over-expressed in a specific subset of mesenchymal cancers (sarcomas). We have used the 293 Flp-In system to establish stably transfected cell lines with constitutive and tetracycline-inducible expression of COAS2. The different cell lines have been used for microarray-based expression profiling experiments, as well as various phenotypic analyses, such as growth in soft agar, chemotherapy sensitivity and tumourigenicity in immunodeficient mice. Several expression constructs have been made for production of recombinant protein in bacteria and insect cells. In addition, attempts to localise the protein and confirm putative interaction partners using MYC-His_6-and EGFP fusion proteins and confocal fluorescence microscopy were performed.

Through collaboration with Dr. Schiene-Fischer's group at The Max Planck Institute in Halle, Germany, the enzymatic activity and possible substrate of COAS2 will be evaluated. We are now establishing different Flp-In systems in a set of mesenchymal cell lines to make an improved model system, and we seek to express COAS2 in a yeast cell model system for monoclonal antibody production.

Abbreviations

A	Adenine
aa	Amino acid
BAC	Bacterial artificial chromosome
BLAST	Basic local alignment search tool
bp	Base pair
BRCA1, -2	Breast cancer genes 1 and 2
BSA	Bovine serum albumin
C	Cytosine
Ci	Curie
C. elegans	Caenorhabitis elegans
CDK	Cyclin-dependent kinase
cDNA	Complementary DNA
CGH	Comparative genomic hybridization
COAS1, *-2, -3*	Chromosome One Amplified Sequence 1, 2, and 3
cps	Counts per seconds
CsA	Cyclosporine A
CsCl	Cesium chloride
Cy3/5	Cyanine dye 3 and 5
Cyp	Cyclophilin
DEAE	Diethylaminoethyl
ddNTP	Dideoxynucleoside triphosphate
dH_2O	Distillated water
dNTP	Deoxynucleoside triphosphate
DEPC	Diethylpyrocarbonate
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
DNR	Det norske radiumhospital HF = the Norwegian Radium Hospital
dsDNA	Double strand DNA
DTT	Dithiotreitol
dUTP	Deoxyuridine triphosphate
E2F	Transcription factor E2F
E.coli	*Escherichia coli*
EDTA	Ethylendiamine tetraacetic acid
(E)GFP	(Enhanced) green fluorescent protein
ERK/MAP kinases	Extracellular signal regulated kinases/mitogen activated protein kinases
EtBr	Ethidiumbromide
FCS	Fetal calf serum
FISH	Fluorescence *in situ* hybridisation
FITC	Fluorescein isothiocyanate
FKBP	FK506 binding protein
FOS	v-fos FBJ murine osteosarcoma viral oncogene homolog
G	Guanine
GAPS	Gamma amino propyl silane
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
HMGIC	High mobility group isoform IC
H. sapiens	Homo sapiens
IPTG	Isopropylthiogalactoside
kb	Kilo base (pair)
LB	Luria-Bertani
MDM2	Murine double minute 2 gene
Min	Minute
mRNA	Messenger RNA
MYB	v-myb myeloblastosis viral oncogene homolog
MYC	v-myc myelocytomatosis viral oncogene homolog
NP-40	Nonidet-P40, e.g. octylphenoxypolyethoxyethanol

OD	Optical density = absorbance (A)
PAC	P1 artificial chromosome
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PNK	Polynucleotide kinase
Pfu	*Pyrococcus furiosus*
PIN1	Protein (peptidyl-prolyl *cis-trans* isomerase) NIMA-interacting 1
PPIase	Peptidyl prolyl isomerase
*RB*1	Gene coding for retinoblastoma protein, pRB
1q	The long arm of human chromosome 1
12q	The long arm of human chromosome 12
rcf	Relative centrifugal force = g
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolution per minute
rRNA	Ribosomal RNA
RT-PCR	Reverse transcription polymerase chain reaction
SDS	Sodium dodecyl sulfate
SSC	Standard saline citrate
ssDNA	Single strand DNA
STD	Standard
T	Thymine
TAE	Tris acetic EDTA buffer
Taq	*Thermus aquaticus*
TBE	Tris boric EDTA buffer
TE	Tris EDTA
TP53	Gene coding for tumour suppressor protein p53
TRIS	Hydroxymethyl aminomethane
tRNA	Transfer RNA
Trp	Tryptophan
U	Units
UV	Ultraviolet
V	Volt
v/v	Volume/volume
WDLPS	Well-differentiated liposarcoma
w/v	Weight/volume
YAC	Yeast artificial chromosome
Ø	Diameter

Human genes should be presented in uppercase *ITALIC*, while proteins should be presented by the same letters in PLAIN uppercase (McAlpine, 1995). These guidelines have been followed for all *H. sapiens* genes and proteins in this work. Genes from other organisms and their gene products have been referred to in lowercase *italic* and lowercase plain, respectively. Some terms, words and methods are presented in *italics*, indicating their description in the **glossary** found in the appendix section A.

1 Introduction

1.1 Cancer in general

Developmental aspects of cancer

The embryo is sectioned into three primary germ layers at the transition between the *blastula* and gastrula stage of mammalian *embryogenesis*. From only 150 cells, all future defined organs and tissues will emerge, leaving a powerful mass of *stem cells* behind as backup for tissue renewal and repair. The blastula consists of an outermost layer of cells, or *trophectoderm,* a fluid-filled cavity called *blastocoel* and a cluster of cells called *the inner cell mass*. Following blastocyst growth and invasion of the uterus epithelium, the inner mass will further be layered into the *ecto*-, *endo*-, and *mesoderm* layers of cells, respectively. Ectoderm is the precursor of skin, nerves and brain. The endodermal cells will differentiate into lungs and digestive organs (kidneys, gut etc), whilst the mesoderm layer of the inner cell mass will be the starting material for connective tissue that support and surround other structures, such as tendons, bone and muscle.

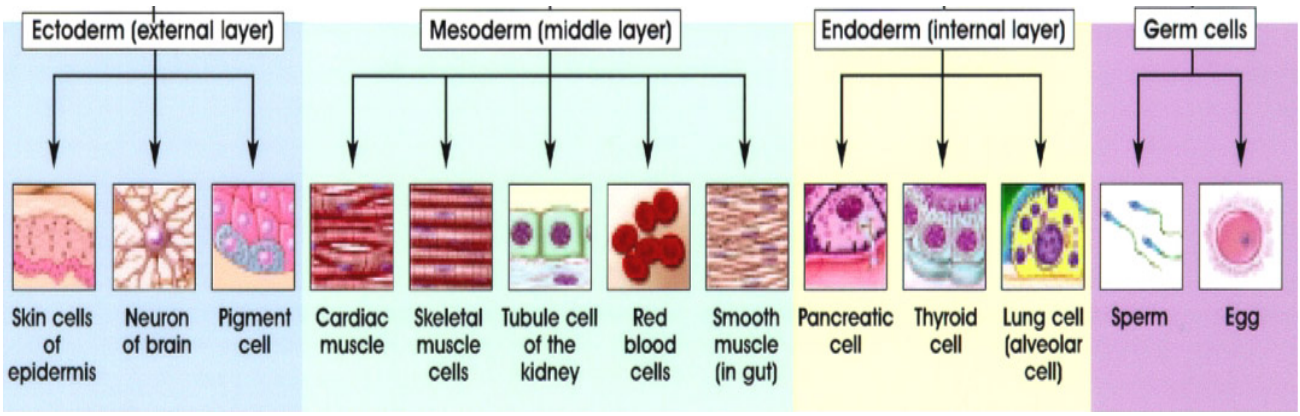


Figure 1.1: Model illustrating the development of cell types from respective embryonal tissues.

This thesis focuses on benign and malignant tumours in tissue of mesenchymal origin. Mesenchymal stem cells (MSCs) originate from the mesenchyme, a loose network of cells within the mesoderm. MSCs can be obtained from bone marrow and cultivated *in vitro*.

Currently, attempts to establish a mesenchymal stem cell line model from bone marrow are made by members of the Myklebost group. Such a model would prove great worth in future cancer research.

The following model modified from (Meltzer, 2001b) depicts the evolution of MSCs and their derivated normal cell types and cancer subtypes:

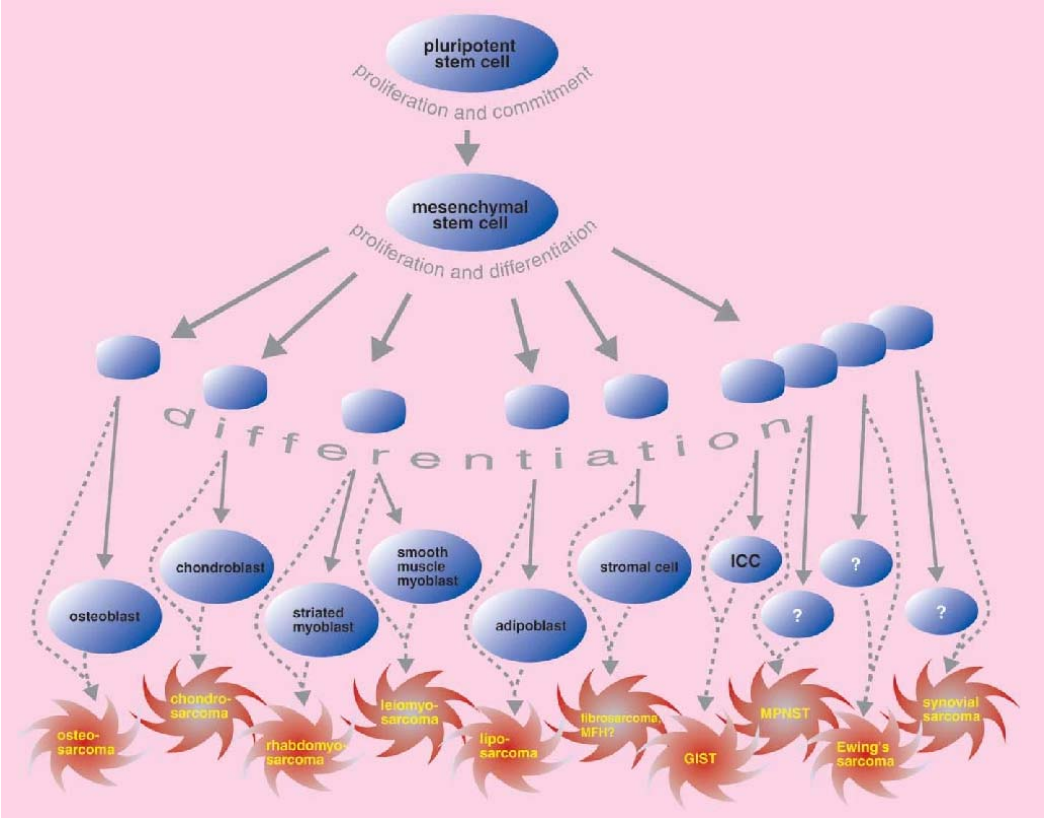


Figure 1.2: Pathways of sarcoma development from tissue of mesenchymal origin (Meltzer, 2001b).

Cancer, as a genetic disease, is driven by mutations in an organism's DNA (Cavenee and White, 1995). There are, however, two key differences between cancer and other genetic diseases. First, cancer is caused mainly by mutations in somatic cells, whereas traditional diseases are caused solely by mutations in the germ line. However, some individuals do have inherited mutations that predispose them to develop specific types of cancer. Second, an individual cancer does not result from a single mutation, but rather from the accumulation of as few as three to perhaps as many as 20 mutations, depending on the type of cancer, in genes that normally regulate cell multiplication. Processes that affect cellular homeostasis may

create a rapidly dividing cell, which may grow to form a cell clone. Whether benign or malign, this cell clone resembles the normal tissue from where it came. Based upon this fact, pathologists generally classify malignant tissues (i.e tumours) as follows: The epithelial tumours from the ectoderm and endoderm are named *carcinomas*, and malignant tumours of mesodermal origin are called *sarcomas* (Enzinger, 1995). There are 300 types of cells or so in the human body, giving more than 200 different types of cancer. With the exclusive exception of leukemias, which grow in the blood as individual cells, most other tumours are solid masses. It is generally agreed that the malignant *neoplastic* cell is less differentiated than the normal adult cell in the organ from which the cancer originates. However, for a long time, there has been a controversy among cancer biologists on the question of whether neoplastic transformation results from de-differentiation of normal adult type cells or from the abnormal differentiation of committed, but not yet mature, stem cells in tissues. It can be concluded, however, that malignant neoplastic transformation occurs only in cells that are capable of dividing (Varmus, 1993; Weinberg, 1996). Currently, the hypothesized *existence of cancer stem cells* holds great promise. Such cells may arise by acquisition of self renewal potential in postmitotic differentiated cells, or mutations in "normal" stem cells. By January 2004, their existence was devised in three cancer subtypes (Pardal et al., 2003). Another peculiar subtype is the *teratomas*, tumours harbouring tissues from three embryonic germ layers. These tumours are most often found in ovary and testis cancer (http://www.kreft.no).

1.2 Cancer and the cell cycle

Most eukaryotic cells live according to an internal program; that is, they go through a series of phases that constitute the *cell cycle*. The processes of cell growth, rest or necessary arrest and possible apoptosis (i.e. programmed cell death) are being tightly regulated at key checkpoints. Monitoring the status of a cell requires the cooperation of many proteins. Nobel Prize Laureate Leland Hartwell illustrated this in the work on the budding yeast *Saccharomyces cerevisae.* The isolation of yeast cell-division cycle *(cdc)* mutants led to the identification of genes that regulate the yeast cell cycle through phosphorylation and protein degradation (Hartwell, 1991). Later, the discovery of yeast protein homologues in *Homo sapiens* affirmed their important role in the eukaryotic cell cycle. In both organisms, the cycle has a mitotic phase (M) where cells divide, and an interphase consisting of a gap phase 1 (G1), a synthesis phase (S) (where the nuclear genomic DNA is duplicated) and a gap phase 2 (G2). In addition

to this, cells can enter a paused, non-dividing, so-called quiescent state (G_0) during G1. Examples of cells maintained in the G_0 state are nerve and striated muscle cells, which do not divide at all (Ford and Pardee, 1999). The complex regulatory events that guide eukaryotic cells from phase to phase will not be fully described here, but some essential proteins should be mentioned: The conserved enzyme family of cyclin-dependent kinases (the CDKs), are activated by binding regulatory proteins called cyclins (Ford and Pardee, 1999). Following activation, the CDKs phosporylate target proteins, acting chiefly on the amino acids serine and threonine (Pardee, 1994);(Pardee, 2003). The CDKs and cyclins A-E with subclasses act in partnership with external growth factors, cytokines and tumour suppressor proteins (Pardee, 1994);(Pardee, 2003). Examples of the latter are E2F, Retinoblastoma (RB) and p53. Together, they survey the cells' condition through tight checkpoint control. Large protein complexes (such as the anaphase promoting complex, or APC) in the G2, M and G1 phase also mark specific inhibitors of cell-cycle events (e.g. p16 and p21), for proteolytic degradation by proteasomes, thus driving the cycle in one direction because of the irreversibility of protein degradation (Ford and Pardee, 1999). As the amount of growth factors and cyclins vary throughout the cell cycle, the cell goes from G1 to S to G2 via M (see figure 1.3). Proteasomal breakdown, cell size and the mentioned external growth factor signals participate in the regulation of cyclin levels (Pardee, 1994). At the restriction point in G1, the cell can proceed to S phase, rest in G1, or go into G_0. Under favourable conditions (if the G1-S checkpoint is passed), cells will enter the S phase. Unfavourable conditions will block or pause the cells in G_0, or hold the cell at the G1-S checkpoint until cyclin D-CDK4 and cyclin D-CDK6 couples have phosphorylated pRB and thereby released the E2F protein, (Nevins, 2001). The transcription of genes needed during S phase was undertaken during G1, as for the G2 genes in mid S phase. Passing through S phase, cyclin A and the respective CDK regulate the active replication complexes that effectively ensure the duplication of DNA. If the genome wide replication is incomplete or if the DNA has been damaged, cyclin A and B with respective CDKs pause the cell at the G2-S checkpoint until necessary steps have been taken. In the M phase, the mitotic spindle separates through a complex regulatory mechanism. The cell cycle maintains the cell number in a living organism, keeping growth and differentation in check.

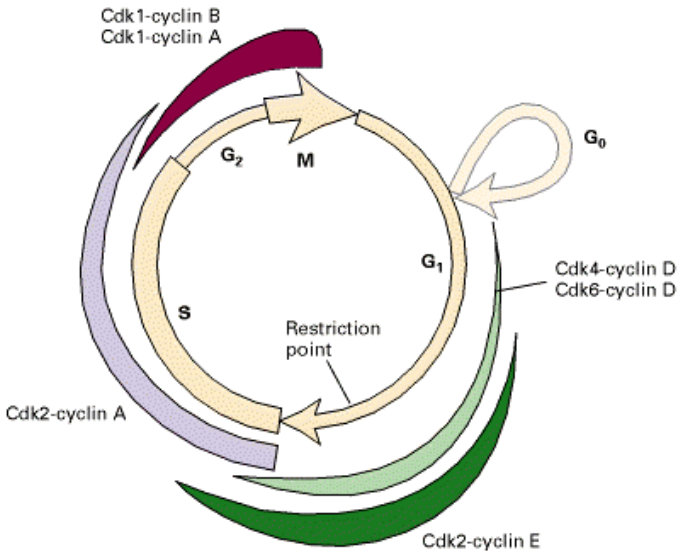


Figure 1.3: CDK-cyclin levels throughout the cell cycle (Lodish, 2000). See text for further information.

1.3 Oncogenes, DNA repair-genes and tumour suppressor genes

Several types of proteins participate in the controlling of cell growth: (Vogelstein and Kinzler, 1993) Growth factors and their receptors, signal-transduction proteins, transcription factors, pro- or anti-apoptotic proteins, cell cycle control proteins and DNA-repair proteins. Of these, mutations in pro- or anti-apoptotic proteins greatly enhance the probability of mutations in other classes (Vogelstein and Kinzler, 1993), whilst mutations in the cell cycle control proteins increase the probability that the mutant cells will become tumour cells (Ford and Pardee, 1999), (Pardee, 1994). In cancer, three main classes of genes are frequently found altered. *Proto-oncogenes* give rise to *oncogenes* through dominant gain of function-mutations, translocations (chimeric gene products or deranged regulation), rearrangements (truncated products) and amplifications (Hayward, 1981; Reddy, 1982; Alitalo and Schwab, 1986; Haluska, 1987). Most proto-oncogene gene products have functions in cell growth or division. The term “oncogenic” essentially means, “will lead to cancer development”, and accounts for the ability to transform either cells in culture or in inducing cancer in animal models. *Tumor-suppressor genes* encode proteins that directly or indirectly perform negative regulation of the cell cycle, e.g. TP53 (Knudson, 1971). These genes are inactivated through recessive loss of function-mutations. Inheritance of a single mutant allele of many tumour-suppressor genes, greatly increases the risk for developing certain types of cancer, e.g. *RB1* (retinoblastoma), *TP53* (Li-Fraumeni syndrome) and *APC* (colon cancer) (Li et al., 1969; Knudson, 1971;

Lindahl, 1996). The third class of genes is involved in DNA repair, the genes belong to one of four subgroups that function either during (mismatch repair, MMR or direct repair) or after synthesis of new DNA (nucleotide excision repair, NER or base-excision repair, BER). Germline mutations in one of these genes may confer an enhanced predisposition to cancer (Berwick and Vineis, 2000; Kohno and Yokota, 2002). Examples of DNA-repair genes linked to cancer in *H. sapiens* include *BRCA1* (breast cancer) and *BRCA2* (breast and ovary cancer) (Smith et al., 2003).

A landmark observation in the search to identify a genetic basis for cancer was reported by Rous in 1911, when he showed that sarcomas could be reproducibly induced in chickens by cell-free filtrates of a sarcoma which had previously arisen in another chicken (Rous, 1911). Sixty years later, the oncogenic region of the Rous sarcoma virus was identified. Further characterization and cloning of the transforming sequences demonstrated that the oncogenicity of the virus was dependent on *v-src*, later found to be a transduced and mutated copy of the c-src cellular proto-oncogene (Varmus et al., 1977). (V = viral whilst the prefix C = cellular in this context.) Subsequently, all oncogenes of acutely transforming RNA tumour viruses have, in fact, been found to be transduced cellular genes, the proto-oncogenes. Although the biochemical mechanisms by which most viral oncogenes cause neoplastic transformation are still not fully defined, the viral oncogenes appear to cause transformation, in general terms, because they are activated versions of cellular proto-oncogenes and/or are expressed aberrantly. In the majority of human cancers, somatic mutations generate oncogenic alleles from proto-oncogenes. The first human oncogene to be identified encodes a constitutively active form of Ras, a signal-transduction protein first isolated from a human bladder carcinoma (Boriack-Sjodin et al., 1998). Activation of a proto-oncogene into an oncogene can occur by point mutation, gene amplification leading to overexpression of the gene product, or gene translocation (Vogelstein and Kinzler, 1993). Examples of common oncogenes in mammalian cancers would be MYC (which retroviral gene is called MC29), MYB (frequently found in myeloblastic leukaemia), RAF and RAS (Alitalo et al., 1984; Vogelstein and Kinzler, 1993; Hesketh, 1997; Dang, 1999; Pinson et al., 2001). All four are among the most frequently altered proto-oncogenes in human cancers (Cavenee and White, 1995; Lengauer et al., 1998). Interestingly, a 2001 review of the findings made by nine groups working in the cell cycle field of research concludes that all oncogenes can deregulate cell growth, but they cannot and will not all lead to cancer if their downstream target proteins are missing (Bartek and Lukas, 2001). In 2001, Yu et al suggested a model of breast cancer

resistance development in cyclin D1 deficient mice, depending on oncogene identity, where the lack of one mutated proto-oncogene conferred resistance to certain types of cancer (Yu et al., 2001). This field of interest is highly promising, and equally enlightening for all types of cancer. Despite the significance of oncogenes in the genesis of many different human tumour types, many of the altered properties of cancer cells appear to be attributable to the inactivation of tumour-suppressor genes (Hesketh, 1997). They encode proteins that delay or inhibit progression through a specific stage of the cell cycle; so called "checkpoint-control" proteins that arrest the cell cycle if DNA is damaged or chromosomes are abnormal. Inherited mutations causing retinoblastoma, an embryonic neoplasm of retinal origin, led to the identification of *RB1*, the first tumour-suppressor gene to be recognized (Knudson, 1971). The G1-S transition in the cell cycle depends upon RB and the E2F protein. RB negatively regulates the transcription factor E2F, but upon RB phosporylation by the CDK2/cyclinE complex, E2F is released. E2F-1 is thus free to activate transcription of important genes needed in the S phase of the cell cycle, (Nevins, 2001). A model of this pathway can be found at: http://www.biocarta.com/pathfiles/h_skp2e2fPathway.asp

The p53 tumour suppressor stands alone as the most extensively studied gene because of its involvement in the genesis of various types of cancer. Bearers of the Li-Fraumeni syndrome hold more than 50 % probability for cancer development in their lifetime, this due to a loss of one allele in the p53 gene (Li et al., 1969). Chromosomal alterations and genomic imbalance activate p53, thus putting the cell cycle on hold until the DNA has been mended. Severe DNA-damage activates p53, which leads the cell to apoptosis (programmed cell death) (Szymanska and Hainaut, 2003). This illustrates the need to inhibit p53 protein for the development of most cancers.

1.4 Genetic basis of cancer

Six necessary steps for neoplastic growth

While researchers have long known cancer to be a genetic disease, it is only recently that they have been able to begin explaining the mechanisms that lead to neoplastic growth. At first, the immense phenotypic and genotypic differences between different types of cancer lead many researchers to treat them as completely distinct diseases. However, persistent research has

slowly revealed an underlying pattern common to all types of cancer. It seems that for a tumour to progress it must overcome a certain number of biological defence mechanisms (Vogelstein and Kinzler, 1993; Cavenee and White, 1995). Cancer thus appears as a multi-step process, where different cancers having different ways of achieving each of these steps. They can be suggested as six major alterations in cell physiology, each of them corresponding to an essential acquired trait: (1) self-sufficiency in growth factors, (2) insensitivity to anti-growth factors, (3) the ability to evade apoptosis and attacks of the immune system, (4) infinite replication potential, (5) sustained angiogenesis and (6) the necessary changes needed to allow metastasis (Cavenee and White, 1995; Klein, 1998). Most often, each of these steps must be successfully taken before a tumour can become truly malignant. This again explains the relatively low occurrence and late onset of cancer during an average lifetime. Turning a normal cell into a malignant cancer cell requires perhaps six specific mutations in the one cell. The chance of a single cell undergoing six independent mutations is negligible, suggesting that cancer should be vanishingly rare. Cancer nevertheless happens because of a combination of two mechanisms:

- Mutations in oncogenes and/or tumour suppressor genes enhance cell proliferation, creating an expanded target population of cells for the next mutation.
- Some alterations affect the stability of the entire genome, at either the DNA or the chromosomal level, increasing the overall mutation rate, (namely translocations, amplifications, deletions and insertions).

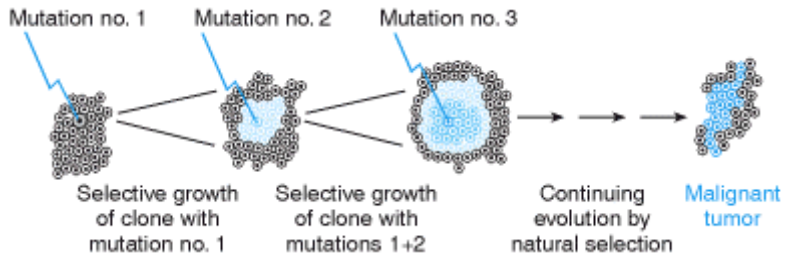


Figure 1.4: Multistage evolution of cancer (Read, 1999). Each successive mutation gives the cell a growth advantage as mentioned in the text.

Accumulating all these mutations nevertheless takes time, so that cancer is mainly a disease of post-reproductive life, when there is little selective pressure to improve the defenses still further. Because cancers depend on some of these mechanisms, they develop in stages, starting with tissue *hyperplasia* or benign growths, while malignant tumour cells have the

ability to invade surrounding tissues or metastasise. Studies suggest that it is not simply the accumulation of mutations, but rather it is their order which determines the propensity for neoplasia, and that only a subset of the genes which can affect cell growth can actually initiate the neoplastic process (Kinzler and Vogelstein, 1996). Recently, the existence of *cancer stem cells (CSCs)* (Pardal et al., 2003), has been given much interest. Their self-renewal potential might largely benefit a tumour's ability to establish and grow from tissue at several stages of development. It is known that the asymmetrical division of normal stem cells give rise to a new stem cell and a progenitor cell (www.nih.gov/news/stemcell/scireport). Hypotheses are that both descendants of a normal stem cell may become a cancer stem cell (Pardal et al., 2003). This may happen either through mutations in normal stem cells, or acquisition of self renewal potential for the partially differentiated progenitor cell. Please address (Pardal, Clarke et al., 2003), for further reading on the subject of CSCs.

1.4.1 Self-sufficiency in growth factors

The ability to produce its own growth factors is the most intuitive of the traits that a developing tumour must acquire, and was therefore the first to be clearly defined. Under normal conditions the cell division cycle is triggered by exogenous growth factors (GFs), thus enabling the cell to pass from its quiescent state (G_0) into an active proliferative phase (G1). However, tumour cells tend to show a very reduced dependence on external factors, and must therefore somehow be able to generate their own growth signals (GS). There are three main ways of achieving this independence: Many *neoplastic* cells acquire the ability to produce GFs to which they are responsive (*autocrine* stimulation). Others modify their external receptors to react to different GFs or over-express them so much that they become hyper-responsive to normal GF levels (Cavenee and White, 1995). The third possibility is to alter the complex molecular cascades that process the signals from the external ligands (Varmus, 1993; Lengauer et al., 1998). The most important of these is *the SOS-RAS-RAF-MAP kinase* pathway, and structurally altered RAS proteins are found in approximately 25% of human cancers (Enzinger, 1995). They continuously release a flux of mitogenic signals into the cell, even in the absence of stimulation from their normal upstream regulators.

1.4.2 Insensitivity to anti-growth factors

In normal tissue, a multitude of anti-GFs help to maintain *homeostasis*. They operate in two ways, either by forcing the cell out of its proliferative state into a quiescent state (G_0), or by further inducing them to enter differentiated postmitotic (following cell division) states. Tumor cells acquire insensitivity to these exogenous factors either by altering their external anti-GF-receptors or by modifying the intracellular processing cascades, but these mechanisms remain not well understood (Varmus, 1993).

1.4.3 Evading apoptosis

Once triggered, apoptosis (programmed cell death) completely destroys the cell within a span of 30-120 minutes. Apoptotic marker genes and signal cascades have been given much attention over the past two decades, and especially the proapoptotic regulator genes. One of the most important of the latter is the p53 tumour suppressor gene, which is deactivated in over 50% of human cancers (Hesketh, 1997).

1.4.4 Infinite replication potential

The three acquired capabilities that have been described till now should, in theory, enable a cell to replicate without limit. However, all mammalian cells except the germ cells seem to carry an autonomous multiplication limit. After a certain number of doublings, a cell enters a final relaxed state named *senescence* and stop growing. However, any extrapolation from cell senescence to aging of the organism is fraught with difficulties (Faragher and Kipling, 1998; Kipling and Faragher, 1999). This can be circumvented through disabling of p53, although the effect has been linked to accumulation of the cyclin-dependent kinase inhibitors $p16^{INK4a}$ and p27 in mice (Malumbres et al., 2000). The cell may then go through several more division cycles before it reaches another state named crisis, due to the progressive decay and ultimate loss of the chromosomal *telomere caps* (Zakian, 1997). This state is characterised by end-to-end fusion of chromosomes, leading to massive karyotypic disarray and cell death (Zakian, 1997). At this stage one occasionally observes the emergence of a cell with the ability to maintain telomeres at a length beyond the critical limit, thus acquiring the ability to multiply without limit (immortalisation) (Zakian, 1997).

1.4.5 Sustained angiogenesis

A growing tumour needs oxygen, and therefore blood vessels, to be able to survive. The tumour cells must co-opt neighbouring endothelial cells into expanding and creating new blood vessels. p53 once again plays an important role. It is known to positively regulate thrombospondin-1, an angiogenesis inhibitor (Yu et al., 1999). The loss of p53 amongst others thus allows the endothelial cells to replicate. Several oncogenes are known to drive angiogenesis, e.g. MYC and *hypoxia*-inducible factor-1, HIFα (Brizel et al., 1996; Maxwell et al., 1997). Though co-operation between different cell types is also observed at other points in cancer development, the need for angiogenesis most clearly illustrates the fact that neoplastic cells must be studied in their physiological environment.

1.4.6 Metastasis

The final, and generally lethal step of cancer development is metastasis, the ability to invade foreign tissues. Successful colonisation of new sites, whether local or distant, requires a capability to adapt to changing tissue environment, and is most regularly achieved through modification of external receptors. However, the tumours of certain cancers become *anaplastic*, which enables them to colonise new tissues with great ease.

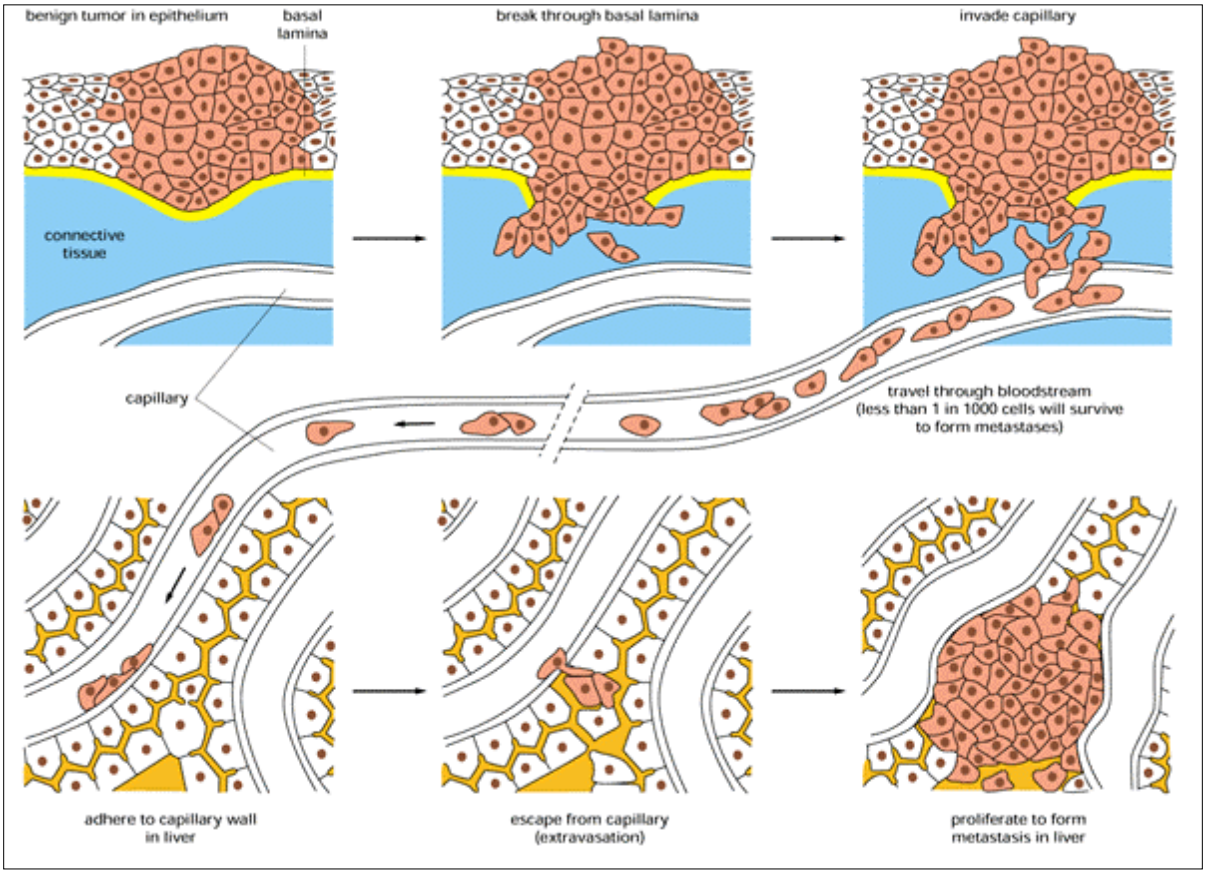


Figure 1.5: Illustration of the spread of a tumour from an organ such as the lung or bladder to the liver (Alberts, 1994). Tumor cells may enter the bloodstream directly by crossing the wall of a blood vessel, as depicted here, or, more commonly, by crossing the wall of a lymphatic vessel that ultimately discharges its contents (lymph) into the bloodstream. Tumor cells that have entered a lymphatic vessel often become trapped in lymph nodes along the way, giving rise to lymph-node metastases..

1.4.7 Genomic instability

A cell must acquire a large proportion of mutations before it becomes truly *neoplastic*. Cancer is said to be the disease of the old, but the high number of single and "independent" mutations are actually highly unlikely to occur at normal rates during a normal human life span. Thus, genomic instability becomes an important enabling characteristic. This is usually achieved through disablement of monitoring systems that normally supervise the genome and repair errors. The most important of these is the p53 signalling pathway, a fact that helps explain why inactivation or mutation of p53 is such an important step in many cancers. Alterations may occur through translocations, deletions, insertions and amplifications. Tumor-type-specific shromosomal rearrangements are thought to be involved in early tumour development

(Rabbitts, 1994), and the unscheduled gene amplification reported to be involved in tumourigenesis (Alitalo and Schwab, 1986).

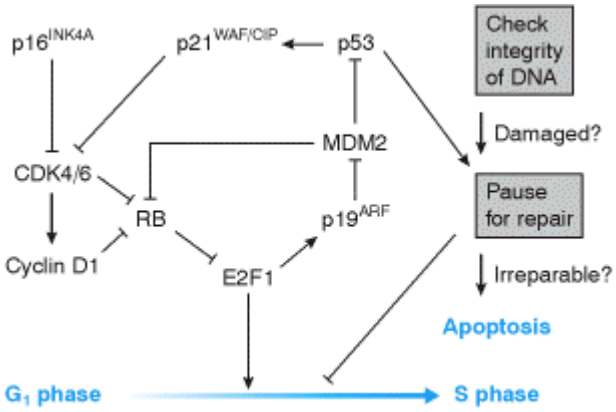


Figure 1.6: Controls on cell cycle progression and genomic integrity mediated by the *RB1*, *TP53, MDM2* and *CDK4/6* gene products(Read, 1999). (RB = retinoblastoma, MDM2 = mouse double minute 2, CDK = cycline dependent kinase.)

Cancer cells possess traits that make them distinguishable from normal cells by microscopic examination. They are usually less well differentitated (a common feature of tumour cells is their failure to differentiate terminally under appropriate conditions either *in vivo* or in culture), have a high nucleus-to-cytoplasm ratio, prominent nucleoli, many mitoses and relatively little specialized structure (Vogelstein and Kinzler, 1993). However, all kinds of degrees exist, making detection at early stages of cancer development difficult.

1.5 Sarcomas in *Homo sapiens*

Introduction

Sarcomas in general are rare malignant tumours, accounting for 1 % of adult neoplasms and less than 10% of pediatric malignancies (Slominski et al., 1999). The word sarcoma is derived from the Greek *sarcos* = fleshy and *oma* = tumour. This huge class of histologically diverse cancers have their origin in mesodermal or ectodermal germ layers, and the tumours harbour traits from muscle, bone, cartilage and other connective tissues of mesenchymal origin. In contrast to carcinomas, which are derived from the endoderm, the sarcomas are not so much defined by their organ of origin as for their molecular pathology, very unlike most epithelial

tumours. The many types of sarcomas do have some common traits, and four familiar cancer syndromes have been associated with sarcomas: Patients with retinoblastoma (RB) germline mutations (Abramson et al., 1984), Li-Fraumeni syndrome including germline mutations of the p53 gene (Li et al., 1969; Malkin et al., 1990), germline loss of the NF1 gene in a malignant peripheral nerve sheath tumour (MPNST) (King et al., 2000), and finally a recently characterised familial gastrointestinal stromal tumour (GIST), associated with germline mutations of the c-kit gene (Nishida et al., 1998). These cancers can be arranged genetically into two broad main categories, although there are traditional classifications in use by both researchers and clinicians (Helman and Meltzer, 2003). According to the pathologists, sarcomas can be described as either soft tissue sarcomas or sarcomas of the bone (Enzinger, 1995). In 2003, Meltzer and Helman suggested a primary category that encompasses tumours with specific translocations seemingly central to their pathogenesis. The second is characterised by complex, chaotic *karyotypes* displaying severe genetic and chromosomal instability.

In addition, pathologists and clinicians often grade sarcomas according to metastatic potential. The histological grade of malignancy, where well-differentiated sarcomas show good prognosis in contrast to the aggressive less differentiated sarcomas, adds to the clinical point of view in terms of classification. Regardless of the classification method in use, the amount of annual sarcoma detections in humans remains relatively stable at 1% of all human cancers. Of these cases, 0,8 % were soft tissue sarcomas, developed from muscle, blood vessels, fat, nerves, deep skin and *synovial* tissues (Slominski et al., 1999). The remaining 0,2 % evolved from bone and cartilage. 156 patients were diagnosed with malignant soft tissue sarcomas and 36 with malignant sarcomas of the bone at the Radium Hospital in Norway during 2000. (22185 Norwegian inhabitants were diagnosed with cancer in 2000, according to the Norwegian institute of population-based cancer research, http://www.kreftregisteret.no). Surgery has so far been the primary approach to treating most sarcomas, assisted by radiation- and chemotherapy in many cases (http://www.kreftregisteret.no). Most sarcomas have abnormalities in signal transduction pathways, exemplified by insulin-like growth factor 1 (IGF-1), retinoblastoma (RB) and p53 signal routes (Helman and Meltzer, 2003). Aside from the translocations and signal transduction pathways mentioned, few initiating events in sarcomas have been described. Amplification of the cell cycle-related genes MDM2 and CDK4, both localised on chromosome 12, occurs sporadically in various sarcomas (Forus et al., 2001a; Helman and Meltzer, 2003). The establishment of xenografts has been proven

difficult in some cases, because of the complexing tumour-specific translocations linked to aggressive growth *in vivo*. Not all grafts have the ability to grow, and there is a minor selection of malignant tumour transplants that grow very slowly as nude mice xenografts, making their use in experiments rather difficult (S. Bruheim, personal communication). Diagnostic procedures may be a field where more research should be made, since as many as 10% of all sarcomas cannot be classified, harbouring very complex karyotypes (Myklebost, 1998). But most importantly, the characterisation of benign growth in lipomas holds potential. The existence of benign mesenchymal tumours, e.g. lipomas, may help researchers understand the transition to the aggressive phenotype held by the malignant counterpart, the liposarcomas.

Table 1.1: Sarcoma subtypes**.**

Type of sarcoma	**Subtype**	**Tissue/cell of origin**
Sarcoma of the bone	Ewing's sarcoma	Bone
	Chondrosarcoma	Cartilage
	Osteosarcoma (OS)	Bone
Soft tissue sarcoma	Fibrosarcoma	*Fibroblasts*
	Leiomyosarcoma	*Smooth muscle*
	Liposarcoma (LS)	Fat
	Malignant fibrous histiocytoma	*Histocytes*/Fibroblasts
	Malignant schwannoma	*Peripheral* nerves
	Rhabdomyosarcoma	*Striated muscle*
	Synovial sarcoma	*Primitive mesenchyme*

1.5.1 Sarcomas of the bone

Bone sarcomas are very infrequent, with osteosarcoma as the most common, followed by chondro- and Ewing's sarcoma, a third variant which occurs mostly in the shafts of long bones. OS appears more often in the bone growth plates, as well as the long bones. Ewing's and OS are found in patients of all ages, although more common in youth, and Ewing's is predominant among the male Caucasian population (Sæter, 1996). Tumors from cartilage-forming chondrocytes appear as chondrosarcomas, a disease with predominance in patients of 50+, affecting the axial skeleton and proximal bones. Metastases often occur early with

deadly outcome as a result of spread to the lung, and the five year overall relative survival rate is approximately 55-70 % for OS and Ewing's and 70 % for chondrosarcoma (Saeter et al., 1997).

1.5.2 Gene amplification and over-expression

Occurring quite commonly in cancer, gene amplification increases dosage of proto-oncogenes that give tumour cells a selective advantage (Schwab, 1999). There are several ways of accomplishing this kind of activation: Coding sequences of the given genes can be found as either direct or inverted repeats, either *in situ* on the original chromosome or dispersed to extra chromosomal markers like homogenously staining regions (HSRs), *double minutes, supernumerary rings or giant rod chromosomes* (Biedler and Spengler, 1976; Kaufman et al., 1979). The latter are characteristic of well-differentiated liposarcomas and some borderline tumour types, often found to be the sole cytogenetic abnormality in these tumours (Dal Cin et al., 1993). In addition to other genetic alterations connected to cancer, such as translocations, deletions or aneuploidy (uneven chromosome number), the increased gene copy number phenomenon is found in a large fraction of human sarcomas. Detection of amplified material has been possible by various molecular and cytogenetic methods; Southern blotting, fluorescent i*n situ* hybridisation (FISH), real time-PCR, comparative genomic hybridisation (CGH) and array CGH. Nils Mandahl's group at the University of Lund (Sweden) described the *COAS2* gene using FISH with *cosmid* probes (Nilsson et al., 2004). *COAS2* was found amplified in lipomatous tumours, with the extra copies located to *ring* or *giant marker chromosomes,* as expected for this type of tumours.

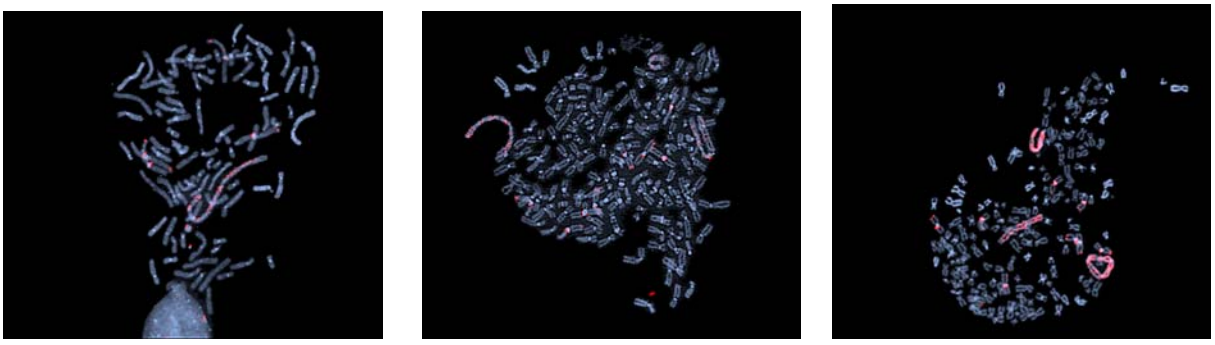


Figure 1.7: Metaphase spread showing additional (red) signals of COAS2 in ring and marker chromosomes, pictures courtesy of Malin Nilsson.

1.5.3 The 1q21-23 and 12q13-15 amplicons

Some regions are highly amplified in a major percentage of human sarcomas, with the 1q21-22 and 12q13-15 *amplicons* frequently seen in tumours of both soft tissue and bone. Forus et al. were the first to identify the important chromosome 1 amplicon in sarcomas (1998), and it has also been observed in breast and ovarian cancer (Kudoh et al., 1999; Tarkkanen et al., 1999).

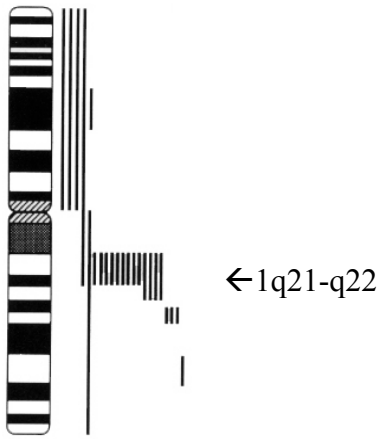


Figure 1.8: CGH studies of chromosome 1 in soft tissue sarcomas. Each vertical line represents amplification in one tumour sample. The arrow indicates the main amplified region, 1q21-q22. The p and q represent the short and the long arm of chromosome 1, respectively. Figure modified after (Forus et al., 1995a).

Amplification in the 1q region is associated with an aggressive phenotype with high metastatic potential, short overall patient survival and poor chemotherapy response (Hirai et al., 1999; Tarkkanen et al., 1999). The *APOA2* gene, and the *COAS* genes are known target genes in this region. Further studies in progress, using a complete tiling path array throughout the region, will give a high-resolution picture of the DNA copy number changes in this amplicon. Other common amplicons in osteosarcomas include the 6p, 8q, 17p11-p12 and the frequent 12q amplicon with amplified *MDM2*, *CDK4* and *HMGIC*. The architectural transcription factor HMGIC is frequently rearranged during the process, and thus probably oncogenically activated (Forus et al., 1993); (Maelandsmo et al., 1995); (Meza-Zepeda et al., 2001). Several other proto-oncogenes are over-expressed in sarcomas. The rapidly increasing identification of marker genes is made possible with microarray techniques. The array technology will be described in further detail in the methods section.

1.6 The immunophilins

Cyclophilin A (Cyp A) was first identified and purified from bovine spleen in 1984 on the basis of its high affinity (K_d =10^{-8}-10^{-9}) for the immunosuppressive drug Cyclosporin A, (CsA) (Walsh et al., 1992). Other groups found similar proteins that catalyzed the interconversion of *cis*- and *trans*-amide bonds adjacent to proline residues in peptide substrates, and termed these enzymes *peptidylprolyl isomerases (PPIase)* (Fischer et al., 1984). Subsequent studies demonstrated that the found PPIase and cyclophilin A were the same molecular entity, and that CsA inhibited its PPIase activity (Fischer et al., 1989; Takahashi et al., 1989). However, the 18 kDa cytosolic protein had an abundant overall distribution which dit not correlate with the selective immunosuppressive action of CsA, and a search for additional members of the **immuno**suppressive drug binding cyclo**philin** protein family commenced. Today, the peptidyl-prolyl isomerase (also known as *rotamase*) family consists of three subclasses based upon drug specificity: The cyclophilins (Cyps), that take CsA, the FK-binding proteins (FKBPs), binding the peptide FK506, and the parvulins, that do not bind immunosuppressants. Ubiquitously expressed, the immunophilins also appear to function as accessory helper enzymes of folding chaperones that provide functional stability to multiprotein macromolecules. Links to axonal transport, synaptic vesicle assembly and the possible role in neuroprotection against normal protein aggregation enhance the image of immunophilins as multitasking proteins (Avramut and Achim, 2003). They are found in various compartments of the cell: the endoplasmatic reticulum: CypB and FKBP13; the nucleus: FKBP25 and 52; and in CypD in mitochondria. CypA and FKBP12 are both cytosolic proteins. The *C. elegans* isoform of CypB has been associated to the gut (Picken et al., 2002). All members share the ability to catalyze the cis-trans isomerisation of peptide bonds N-terminal to proline residues in polypeptide chains; a rate-limiting step in protein folding and regulation of ligand selection (Brazin et al., 2002; Shaw, 2002). The PPIase activity is also essential for skeletal and muscular differentiation (Hong et al., 2002). Parvulin activity is directed by the prior phosphorylation of target proteins in both yeast and mammalian cells (this is also the case for some cyclophilins), and they are structurally distinct from FKBPs and cyclophilins. Cyclophilins have an eight-stranded β-barrel that forms a large hydrophobic pocket located at the surface of the central core in which CsA binds (Carpentier et al., 2002), whereas FKBPs consist of an amphiphatic (i.e. both hydrophilic and hydrophobic) five-stranded β-sheet that wraps around a single, short α-helix (Michnick et al.,

1991). Parvulin family members possess a PPI domain of a half β-barrel, its four antiparallel strands surrounded by four α-helices.

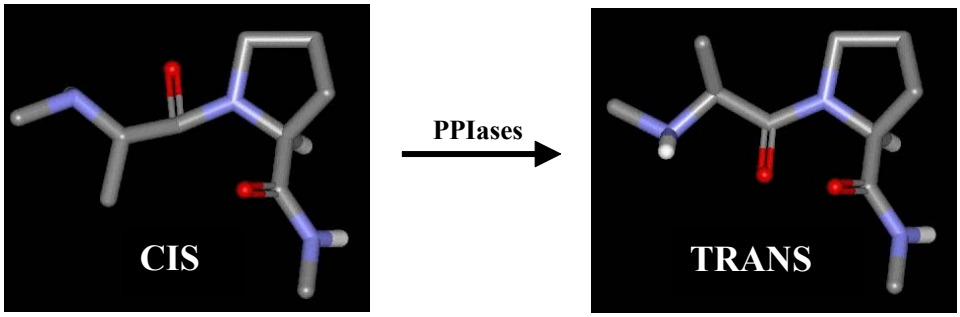


Figure 1.9: Peptidyl-prolyl isomerization of the amide bond preceding Proline. This bond has partial double bond character and can exist as both *cis* and *trans* isomers. For all amino acids except proline, the *trans* isomer is the preferred conformer because of steric hindrance in the *cis* form, whereas the peptide bond preceding proline residues are equally stable as both *cis* and *trans* isomers. PPIases have the ability to catalyze the interconversion between the *cis* and the *trans* isomers. See text for further information.

A peptide bond has partial double bond character and can exist as either cis or trans isomers. For all amino acids except proline, the *trans* isomer is the preferred conformer because of steric clashes in the *cis* form. In contrast to this, the peptide bonds preceding proline residues are equally stable in both isomeric states (Arevalo-Rodriguez et al., 2000). In the living organism, cis-trans forms of peptidyl-proline bonds exist in a 5% - 95% equilibrium (Reimer et al., 1998). In the cell, these bonds are thought to be catalyzed by the ribosomes as trans-isomers, and spontaneous OR assisted (by either protein disulfide isomerases or PPIases), isomerisation yields the cis form (Fischer and Schmid, 1990).
The phenomenon of peptidyl-prolyl isomerisation of Xaa-Pro amide bonds (where Xaa is any amino acid) has been extensively studied. The parvulins have high specificity for phosphorylated pSer/Thr(PO_3H_2)-Pro motifs, and a few Cyps use this well-studied substrate (Ryo et al., 2001; Messenger et al., 2002). Concerning the immunosuppressant-binding PPIs, X-ray and NMR structures of the CypA/CsA and FKBp/FK506 complexes suggest that these proteins promote the cis-trans isomerisation of an Xaa-Pro bond by preferentially binding it in a twisted (nonplanar) conformation. They thus follow a transition state binding type of catalysis, whereof some substrates may be phosporylated as an initiating signal. The early step of conversion from cis- to trans-X-Pro bonds is thought to release the substrate from the enzyme for further processing or post-folding modification. The phosphorylation dependent PP isomerisation has also been suggested to act as a signal enhancer for kinase activity, where

phosphorylation followed by cis-trans conversion tags target proteins for ubiquitin-mediated proteasomal degradation (Hsu et al., 2001). A model for PPI inhibition has been suggested where the immunosuppressant drugs act as substrates where this cis-trans conversion does happen, but not to its full extent. This resulting in a permanently bound substrate with both cis-and trans-structures accomodated in the same active site with minimal changes in the path of the polypeptide, resulting in occupancy of the pocket and full inhibition of the PPI activity (Howard et al., 2003). Various amino acids are responsible for the PPI activity in the respective enzymes, which again helps explain the difference in drug specificity. As an example, steric hindrances in the natural state of many X-Pro bonds and Proline side chain oxygen binding to Arg55 in CypA add up to give this particular cyclophilin a preference for Gly-Pro-bonds (Howard et al., 2003). However, the preference is but that, and Cyp catalysis of cis-trans conversion does not depend on the identity of the preceding residue X, (in XC(=0)-Npro bond rotation), but that of FKBP does (Hur and Bruice, 2002). Surrounding domains account for interfamiliar differences and specificities of the respective proteins. WW-domains in parvulins connect them to transcription factors, protein kinases (Messenger et al., 2002; Shaw, 2002), and Tetratricopeptide repeat domain (TPR) immunophilins are linked to steroid receptor signalling through binding to heat-shock chaperone Hsp90; FKBP51,FKBP52 and Cyp40 (Ratajczak et al., 2003). There are glycosaminoglycan (GAG)-binding domains in CypA and B, organelle-signaling sequences (in CypB+D), and RNA Recognition Motifs, with 1 RRM in hCyp33 and in the *D. melanogaster* orthologue (Anderson et al., 2002). Through these domains, the rotamases attach to interaction partners and possible substrate(s). Some have secretory signals, such as CypD. These peptides enable the sorting of proteins from the endoplasmatic reticulum (ER) to the cell membrane. CypB has such a 25-amino acid signal at the N-terminus, and is secreted during immunogenic responses to a final concentration of <5nM, whilst amino acids 1-29 of CypD is a mitochondrial target sequence, cleaved after import into mitochondria. CypD is specifically up-regulated in human tumours of the breast, ovary and uterus, and has been identified as a component of the permeability transition pore in mitochondria (Schubert and Grimm, 2004). Both Cyp A and B bind to T cells via binding to the glycosaminoglycan (GAG) chains of heparins (Allain et al., 2002; Yurchenko et al., 2002). In addition, CD147/EMMPRIN, a highly glycosylated cell surface protein of the immunoglobulin superfamily, has been shown to act as a costimulatory molecule in cyclophilin-mediated signalling events, in addition to its role as signaling receptor for CypA (Carpentier et al., 2002; Yurchenko et al., 2002).

The work presented by Carpentier et al., (2002), suggested a scenario where CypB is immobilized on the T-lymphocyte surface bound to proteoglycans, then isomerises an exposed X-Pro bond in a functional receptor and thereby induces the biological response and ultimately the T cell's adhesion to the extracellular matrix via fibronectin. Importantly, this effect appears to be targeted predominantly to memory $CD4^+$ cells, suggesting a role for CypB, recruiting T cells to infected tissue in vivo (Carpentier et al., 2002). De Ceuninck elucidated the Cyclophilin-GAG-interaction further in 2003, by connecting CypB to chondrocytic heparan sulfate proteoglycans (HSPGs) and matrix metalloproteinases, MMPs. Earlier studies have shown that the attachment of HIV-1 to target cells is a multi-step process that requires an initial CypA-heparin interaction (Sherry et al., 1998; Saphire et al., 1999). Other immunophilins have been linked to cancer, such as the parvulin Pin1, a cell cycle regulator (Basu et al., 2002; Wulf et al., 2002). Over the last decade, several studies on immunophilins in yeast have found that all but one parvulin null mutation are not lethal, indicating that the remaining yeast immunophilins are not essential to neither growth nor cell survival. Intriguing, however, were studies where the Pin1 yeast homologue *ess1* null mutants were shown to have an increased sensitivity to CsA. This was supported by equally interesting crosstalk and complementation studies of *ess1/cypa* across subfamiliar boundaries in yeast (Arevalo-Rodriguez et al., 2000; Fujimori et al., 2001). The authors concluded that *cypa* becomes essential when *ess1* function is compromised. Complementation as a cancer development/progression inhibitory mechanism illustrates fully the potential for immunophilins in cancer research.

1.6.1 Natural substrates of immunophilins

Little is known about immunophilins and their substrates in vivo. Receptor kinases, calcium release channels and steroid receptor complexes have been linked to FKBP12, FKBP12, Cyp40, FKBP51 and FLBP52, respectively. The Cyclophilin H interacts specifically and stably with the U4/U6-specific protein 60K involved in splicing. This through a second protein-protein interaction site, named a "CypH like cleft" formed by the α1-β3 loop (Ingelfinger et al., 2003; Reidt et al., 2003). Otherwise, most of the interactions described here are transient only, except when the cyclophilins posess a TPR domain or a Ran-binding domain (Ferreira et al., 1996). Transitional interaction is a keyword for immunophilins in general, as they perform their rotamase activity without being permanently coupled to the

substrate. Murine Cyclophilin C has an alleged 77 kDa membrane protein substrate named Cyclophilin C associated protein (CyCAP) (Friedman and Weissman, 1991). CypA and FKBP12 interacts with YY1, a zinc finger transcription factor (Yang et al., 1995). The nonreceptor protein tyrosine kinase Itk alternates between the CypA bound monomer and the phospholigand bound mono- or dimeric state. CypA effectively regulates Itk through transition state binding, switching the kinase from an inactive cis-state to the active trans-isomer. More importantly, the interaction of the retinoblastoma gene product RB and Itk links CypA closer to T-cell activation and regulation (Cui et al., 2002).

1.6.2 The pseudosubstrates of immunophilins

As for CypA, given the NMR evidence indicating that CypA-bound CsA undergoes a cis to trans-isomerization, the drug can be thought to be a slow binding pseudosubstrate (Fesik et al., 1990). A parallel occurs in FK506 bound to FKBP (Van Duyne et al., 1991). The difference of bound and soluble CsA is most notable around the 9-10 amide bond, which is cis in solution but trans when bound to the protein. Therefore, each immunosuppressant drug may inhibit its respective immunophilin by selective accumulation as a *trans* product analog, selectively stabilized in the active site and slowly released (Walsh et al., 1992). Several CsA analogues lacking the immunosuppressiveactivity have been developed. These are of interest in the field of anti-HIV therapy, where the prevention of CypA binding to the HIV1 gag protein may prevent replication of the HIV virion (Franke and Luban, 1996).

1.6.3 The cyclophilin class of Peptidylprolyl Isomerases

Cyclophilins are highly conserved genes from prokaryotes to eukaryotes and archaea, and they all share a 109 amino acid-domain surrounded by unique domains involved in organelle and domain targeting (Walsh et al., 1992). Several subgroups are known: at least 12 unique proteins are found in humans (Galat, 2003). Of these, crystal structures of 11 isoforms can be found in structure databases in July 2003 (Galat, 2003). The most common types have all a high extent of amino acid similarity to the family prototype CypA (an example: *COAS2:* 84%, *CYPB*: 65%). The proteins have multiple roles in vivo, targeted to pre-mRNA splicing (Horowitz et al., 2002), transcription factor regulation (Leverson and Ness, 1998) and peroxiredoxin activation (Lee et al., 2001). The cyclophilin family includes CypA, B and C-

like proteins, divergent loop cyclophilins and multidomain proteins (Dornan et al., 2003). CypA has several pseudogenes (www.pseudogene.org). Regardless of amount or false positives, they have a central core composed of eight antiparallel β-strands that form a right-handed β-barrel. The barrel is overlayed by connecting loops and α-helices with additional β-sheets at the termini. The 11 residues responsible for PPIA activity and substrate/CsA binding lie in close vicinity within a large hydrophobic pocket at the surface of the central core (Dornan et al., 2003). In line with differential interaction with calcineurin, different residues present in two loops and in the 3_{10} helix residing near the CsA binding site are involved and altered from cyclophilin to cyclophilin (Etzkorn et al., 1994). Studies of cyclophilin-ligand complexes show that the substrate proline always adopts a cis-conformation, binding via Arg55 and Asn102.

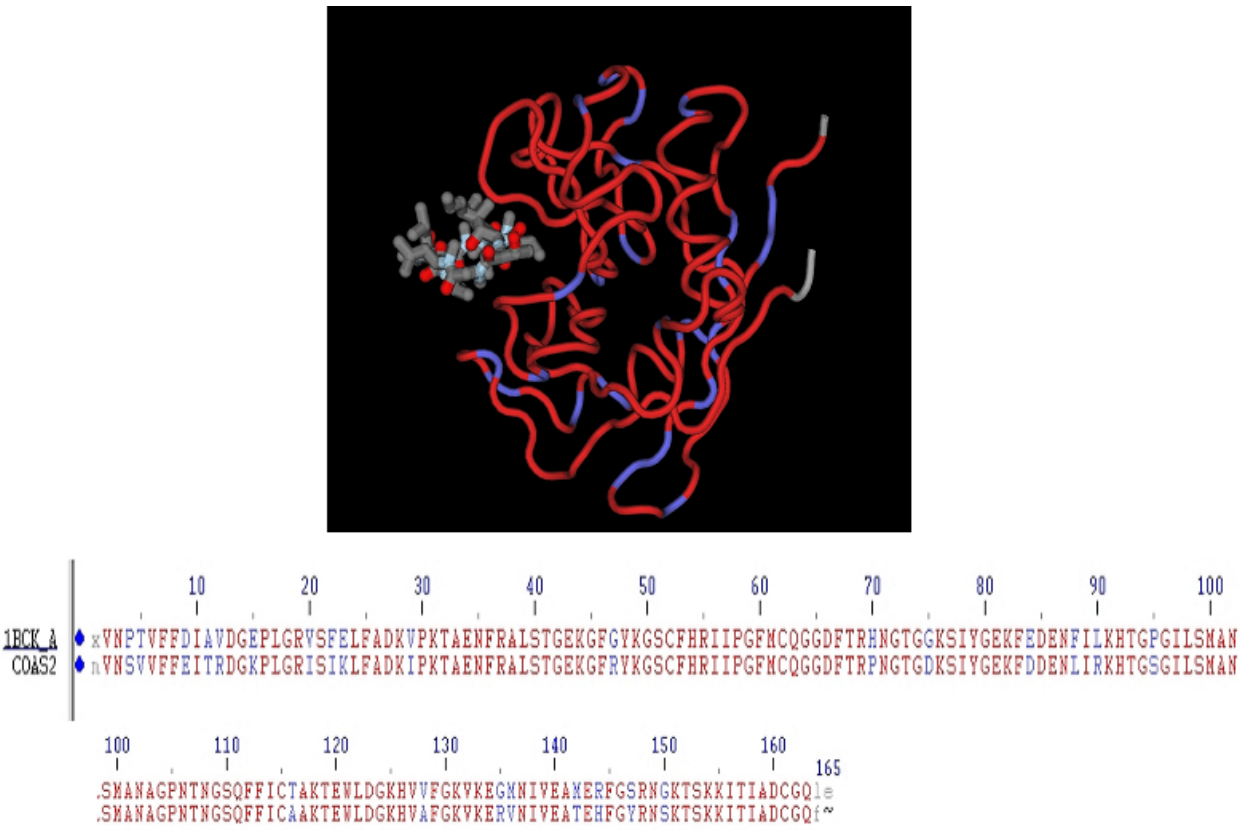


Figure 1.10: The proposed structure of COAS2 with bound cyclosporin A, as wieved with the Cn3D software (NCBI). Blue = amino acids differing from CypA, Red = amino acids identical to CypA, Grey = N- and C-termini. The colors are kept in the sequence aligment of human cyclophilin A and the translated COAS2 mRNA, as shown below. Figure courtesy of Ms B. Lygren.

The basic C-terminal **148RRKE155 domain** in CypA and the respective 3KKK5 plus 14YFD16 N-terminal domains in CypB have been assigned to mediate GAG-binding through an ionic, temperature-dependent interaction (Carpentier et al., 2002). The interaction is very likely to be of transient nature, since most cyclophilins bind to ubiquitously expressed ligands except for the stably interacting CypH-orthologues and perhaps some of the RNA-b binding cyclophilins (Shaw, 2002). Recent publications have reported a degree of cis-trans isomerisation in HIV-virions during interactions between CypA and the HIV Vpr N-terminus (Bruns et al., 2003). Figure 1.11 illustrates a possible model for this CypB-GAG/ CypA-Vpr protein interaction.

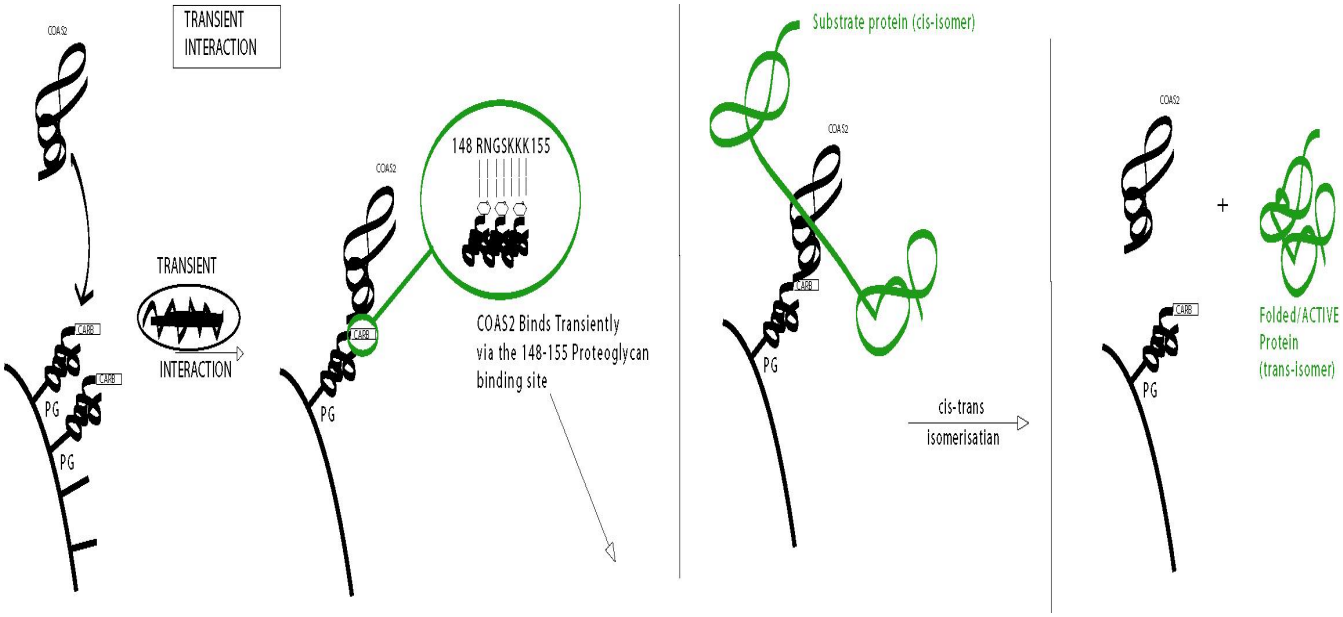


Figure 1.11: Possible model for transient cyclophilin B-binding to GAGs followed by cis-trans isomerisation on a third molecule. In a second model proposed by Carpentier et al., (2002), the third molecule could be a cell surface receptor.

With a huge number of CypA-paralogues seemingly non-essential for life in the host organism, one can hypothesise about the targets and functional aspects of small cyclophilins in vivo. In this work, *COAS2*, a *H. sapiens* class A cyclophilin with 84 % amino acid identity to CypA has been studied at the DNA, RNA and protein level.

1.6.4 Immunosuppressive agents bind immunophilins

Cyclosporin A, FK506 and rapamycin

Cyclosporin A, or CsA, is a cyclic undecapeptide derived from a Norwegian soil fungus; *Tolypocladium inflatum*. CsA has revolutionized the field of transplantation medicine since its discovery at Hardangervidda, and was characterised in the mid-1970s (Borel et al., 1976). Another immunosuppressant is FK506 or tacrolimus; a macrolide compound from the filamentous Japanese bacterium *Streptomyces tsukabaensis*. A third example and *Streptomyces* macrolide is rapamycin or sirolimus from *S. hygroscopicus*. Rapamycin was found at the Easter Island in the early 1970s and evaluated initially for antifungal activity (Sehgal et al., 1975). All three exert their pharmacological effects through the binding of immunophilins: CsA to CypA, FK506 to the FK506 binding protein (FK506P), and rapamycin to other FKBPs. The binding blocks the intrinsic PPI-activity of the immunophilins, but blocked immunophilins do not directly inhibit the immune response (see next paragraph). Recently, another immunosuppressant drug that shares its intracellular target with CsA was discovered. Named sanglifehrin A, it exerts its activity in yeast cyclophilins via the same specific tryptophan; number 128 (Pemberton and Kay, 2003).Both the immunophilins and their molecular targets are found in many cell types, and therefore have effects on many types of tissues. CsA is also linked to muscle differentational arrest: In a FASEB Journal express article (Hong et al., 2002), the authors showed that CsA blocks muscle differentation by inducing oxidative stress and by inhibiting the PPI activity of CypA. The PPI-activity is necessary for muscle cell differentiation and thus protects myoblasts from CsA-induced cytotoxicity. In addition, CsA has higher affinity for Cyclophilin B and C than for CypA (Carpentier et al., 2002), and it may be likely that CsA mediates the above effect by binding to more than one type of cyclophilin. Both CsA and FK506 are efficacious for the treatment of autoimmune and inflammatory diseases, such as psoriasis, Behcet's disease and insulin-dependent type-1 diabetes (Matsuda and Koyasu, 2003).

1.6.5 Immunosuppressive agents exert their effect through a third molecule

Calcineurin (CN), or protein phosphatase 2B, was originally identified as a novel protein in neural tissue which bound Ca^{2+} and inhibited the cAMP phosphodiesterase activator calmodulin (Klee et al., 1979). Subsequently, it was found to be a calcium- and calmodulin-binding protein serine/threonine phosphatase (Aramburu et al., 2000). Conserved in all eukaryotes from yeast to humans, it is in the inactive state composed of two subunits (Fox and Heitman, 2002). One catalytic subunit, calcineurin A, and a regulatory subunit, calcineurin B, with molecular weigths of about 58 to 69 kDa and 16 to 19 kDa, respectively. Upon binding of T-cells to T-cell receptors on damaged cells, CN is activated as the intracellular calcium ion levels rise. Ca^{2+} -calmodulin will free the active site upon binding of the catalytic subunit A. Two Calcineurin B-binding proteins have been identified, namely tubulin and heat shock protein 60 (Li and Handschumacher, 2002). Calcineurin exerts biological functions in higher eucaryotes, inlcuding sperm motility, ion channel regulation, muscle glycogen metabolism, *cardiac hypertrophy T-lymphocyte activation and neurite* extension (Li and Handschumacher, 2002). It has also been found to dephosporylate a number of phosphoproteins including histones, myosin light chain (gene symbol MYL) and the regulatory subunit of cAMP-dependent protein kinase (Aramburu et al., 2000). In their paper from 2002, Neal and Clipstone show that CN mediates the Ca-dependent inhibition of adipocyte differentiation in 3T3-L1 cells. Describing calcineurin's role as a regulator of adipogenesis, the link to CsA has been given proper attention. Their data suggest that CsA and FK506 treatment leads to obesity, and this may partly explain why increased obesity, *hyperlipidemia* and type II-diabetes have been reported in patients treated with these drugs (Neal and Clipstone, 2002). Lymphocyte activation requires the dephosphorylation and revealing of a nuclear transport signal of the Nuclear Factors of Activated T cells (NFATs). This important checkpoint is regulated by calcineurin. Upon addition of CsA or FK506, the drug-immunophilin complex, and not the immunophilin alone will bind to and sequester CN and thus the activation of an immunogenic response. As a result of the drug-immunophilin-blockade, the NFAT class of transcription factors will not be able to translocate into the cell nucleus, and the necessary lymphokines and activating mediators of the immune reaction will not be transcribed. (Interleukin 2 (IL-2), whose synthesis by T lymphocytes is an important growth signal for T cells, Interleukin 4 (IL-4), Interpheron alpha (IFNγ), the CD40 ligand etc.) Calcineurin exists at lower levels in T cells compared to other types of cells, and is therefore more sensitive to drug inhibition (Aramburu et al., 2000).

Recently, Cyclosporin's mechanism of action have been thoroughly discussed, something which led Satoshi Matsuda and Shigeo Koyasu in 2003 to suggest that CsA and FK506 exert their immunosuppressive activity not only through binding immunophilins and Calcineurin, but also by the suppression of Jun Kinase (JNK) and p38 pathways. However, the potential role of immunophilins in this process has not yet been fully elucidated. Unlike CsA and FK506, Rapamycin inhibits by blocking downstream proteins in the signal transduction pathway triggered by ligation of the IL-2 receptor (IL2R). A rapamycin-FKBP12 complex binds to the protein kinase Mammalian target of rapamycin (mTOR), and effectively inhibits further phosphorylation of intracellular targets. mTOR is a member of the PI3K kinase family based on its homology within its catalytic domain. Because rapamycin has different pharmacological activities from CsA and tacrolimus, trials are being undertaken to see if a combination of the drugs might provide more effective and safer treatment with less side effects. For additional information, please see Nature Insight reviews Vol 411 May 2001.

1.6.6 CN inhibition by immunosuppressive drugs

The X-ray structures of the ternary CN-FK506-FKBP and CN-CsA-CYPA complexes revealed in molecular detail how these natural products achieve specific inhibition of calcineurin. First, both inhibitor complexes form extensive drug-protein and protein-protein contacts with CN. Residues 3-9 of CsA, the presumed "effector domain", create a hydrophobic interface with Cn, containing two buried hydrogen bonds (Jin and Harrison, 2002). Second, both inhibitor complexes bind to the hydrophobic interface between the subunits in site that is unique to CN. Third, the active site of CN is not occupied in either complex, but one assumes that the inhibition occurs by occlusion of large protein substrates by the inhibitor complexes. Most importantly, the sites at which the two drug-immunophilin complexes bind to CN are overlapping but distinct, providing a molecular explanation of how two completely dissimilar complexes can both inhibit the same target in a competitive fashion (Fox and Heitman, 2002). In their excellent JBC paper, Yukun Cui et al. (2002) investigated the interaction of the retinoblastoma gene product with CypA. According to the authors, RB can possibly prevent CsA-inhibited NFAT-signalling (Cui et al., 2002).

1.7 Background

The highly amplified 1q21-22 region was characterised in a YAC-based sarcoma panel by Forus, Meza-Zepeda and colleagues at the Norwegian Radium Hospital using comparative genome hybridisation (CGH), and fluorescent *in situ* hybridisation (FISH) (Forus et al., 1995b; Forus et al., 1998). Early CGH results were confirmed by FISH, where the yeast artificial chromosome (YAC) CEPH789f2 was chosen to select cDNAs corresponding to target genes (see Figure 1.12). A linked cDNA library from an osteosarcoma cell line with high-level amplification of the designated area was constructed and hybridised to biotinylated YAC sequences.

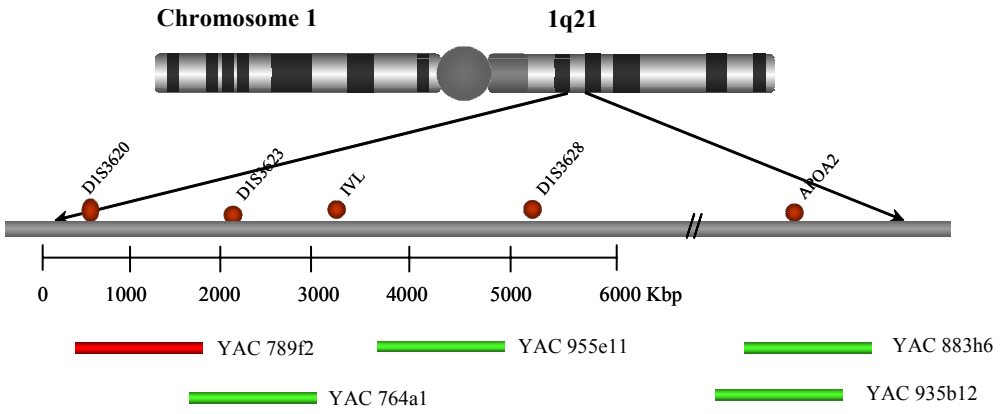


Figure 1.12: YACs covering the 1q21-23 region. The different YACs containing segments of human 1q21-23 DNA used to find the amplification status of 1q21 in a panel of sarcomas are shown in green, and the YAC 789f2 most frequently detecting 1q21 high level amplification, shown in red, was chosen to select cDNAs corresponding to genes in this YAC. The two YACs at the right cover the 1q23 region. Red circles represent different markers in the region of interest. Figure kindly provided by Dr. Meza-Zepeda.

After collection of biotinylated cDNAs with streptavidin-coated magnetic beads and testing for false positives, three novel coding sequences remained as positive hits. They were named Chromosome One Amplified Sequence 1, 2 and 3, (*COAS1, -2* and *-3)* (Meza-Zepeda et al., 2002).

- *COAS1* shows no homology to any known protein families. The 10 kb transcript apparently consists of numerous repeats, yielding a pattern of repeats within repeats, and the function of *COAS1* is not known.

- *COAS3* had some sequence resemblance to the Profilin 1 gene, but the eventual *COAS3* gene product has not been submitted to any further investigation, because it is less amplified than the others.

COAS2, however, was stated a novel member of the cyclophilin family of proteins, with 84% amino acid identity to CypA. It has been reported to be over-expressed in liver cancer by S. Tanaka, (directly submitted to the NCBI database in May 2002, not published), and exists as *cyclophilin-LC* with NCBI accession number AB084917. The gene has multiple allelic versions. The one exon (495 bp) mRNA allegedly codes for a protein of 164 amino acids. The overview article by A. Galat (2003) points out the major problem concerning the characterisation of this gene. Para- and orthologues are very similar, and numerous allelic versions exist, as well as non-translated mRNAs of cyclophilin pseudogenes. *COAS2* was given the LocusLink-number 164022, it bears all signs of a protein-encoding gene, but has until now only been identified as mRNA *in vivo* (Meza-Zepeda et al., 2002).

Earlier work on *COAS2* using transient transfection of egfp fusion constructs gave egfp-COAS2 a speckled distribution pattern within the transfected fibroblast cell, but only very few cells showed fluorescence compared to egfp controls, suggesting some kind of experimental artifact (Lygren, 2002). Interaction partners were searched for by yeast two-hybrid analysis, where the Filamin A and C genes plus the proteasome subunit 7 alpha gene were identified as positives. Localization studies attempting to assay the filamin-COAS2 link did not come to any final conclusion, since the commercial anti-chicken filamin antibody used still has not been fully tested for human filamin specificity. This work set out to confirm and complete some of the early findings.

1.8 Aims of this study: Functional studies of COAS2

Ms Birgitte Lygren performed several experiments with *COAS2* during her time as a MSc-student; some have been referred to and even reprinted here after kind permission. Dr. Leonardo A. Meza-Zepeda is also a major contributor, and their collaborate work have been of great importance to this study. Without the aid of Mr. David J. Warren and Mr. Jørn Henriksen, the protein analysis would never have been made possible.

Initially, the aims were clearly defined, encircling the expression analysis of *COAS2* using cell line material, northern hybridisations and cDNA microarrays, in addition to protein

expression for monoclonal antibody production. The primary goal was nevertheless to establish stably transfected mammalian cell lines as the starting material for the entire microarray-study. We also hoped for the microarray-based transcription profile to link COAS2 to gene regulatory pathways. Characterisation of the mammalian cell clones, confirming the proper insertion, expression and cell cycle regulatory effect of the gene and its product was thus prioritised. Later, the project evolved to include confocal microscopy-based localisation studies using Flp-In 293 cell lines stably transfected with MYC-HIS_6-tagged COAS2. We finally transfected both *egfp-COAS2* and *COAS2-egfp* constructs into the Flp-In system to localize the recombinant proteins with confocal microscopy, aiming to confirm Birgitte Lygren's results achieved with transient transfection systems.

2 Methods, short protocols and related theory

Each topic and some related basic theory is presented in brief, followed by the protocol. In some cases, the protocol included from the supplier was used, and eventual minor changes to the commercially available protocols are indicated. Plasmid maps, list of reagents, solution recipes and a list of equipment used in this work can be found in the appendix, whilst flow charts for the cloning procedures are located in the results section.

2.1 Some standard methods used in molecular biology

2.1.1 Growth and transformation of bacteria

Escherichia coli strains used in this work were all grown in Luria-Bertani (LB) medium supplemented with the appropriate antibiotic (Sambrook, 2001). After 12-18 hours of growth, a freezer stock was prepared with 1 ml of the total culture volume to 25-30% (v/v) of glycerol, and stored at -80°C (Sambrook, 2001). The remaining culture volume was harvested by a centrifugal spin and the pellet was stored at -20°C, awaiting plasmid isolation. *A plasmid is a small, circular double stranded fragment of DNA, harbouring an antibiotic resistance marker gene. In vivo*, plasmids are located in bacterial cells, where their natural role is not well defined. But plasmids do express proteins conferring antibiotic resistance to the host cell, and its autonomic replication in bacteria inspired researchers to isolate and use plasmids as DNA transfer vectors. This technology is well established, and commercially available vectors of many types and sizes are available. Using restriction enzyme techniques (see 2.1.4) the user can insert a gene of interest into the polylinker of a chosen plasmid. Engineered bacterial strains are often preferred hosts, into which the constructs are inserted by transformation (see further on). When grown in a shaker for 12-18 hours, a bacterial culture thus creates a large amount of plasmid DNA, which can be purified using standardized methods. Commercial bacterial strains were all *transformed* using SOC-medium and modified heat-shock methods, as described in (Sambrook, 2001). Basically, 30-50 µl of a chosen competent bacterial suspension, (e.g. JM109, DH5α or DH10α, as used in this work), is thawed on ice, quickly mixed with 1-50 ng of the plasmid to be transferred, and the mix is submitted to a cycle of cooling on ice, heating at 42°C for 30-45 seconds followed by re-

cooling on ice. This heat-shock method will permeabilize the bacterial cell membrane and hence permit the constructs to enter. After 2-3 minutes on ice, different dilutions were plated on LB plates with selection marker antibiotic, and incubated at 37°C overnight. Positive clones were picked using a sterile pipette tip, and grown in LB for 12-15 hours in a shaker at 37°C to enable harvest, storage and further isolation of the plasmid. A modified strain-specific heat shock transformation protocol can be found at the website of the supplier.

2.1.2 Isolation of plasmid DNA from bacterial cultures

1-5 ml of bacterial culture is the starting point for a plasmid "miniprep", whereas the amount increases to 100-250 ml for a "midiprep" and approximately 500 ml for a "maxiprep". The plasmid miniprep protocol used in this study is based upon alkaline lysozyme/detergent lysis of cells as described by H.C.Birnboim in 1983. After a lysis period of 5 minutes maximum, a swift gentle neutralisation will prevent bacterial genomic DNA from being sheared. Now, the high salt concentration causes cell debris, genomic DNA, total RNA and proteins to aggregate and be pelleted by the following centrifugation. The small plasmids will quickly renaturate and remain in the supernatant. In the commercial miniprep kits used in this work, the supernatant is loaded onto a spin column having a plasmid-binding silica membrane. Repeated ethanol washes remove most contaminants, and following elution steps with alkaline 10 mM Tris solutions yields relatively pure plasmid DNA. Along with increasing starting material comes the need for larger columns, which makes spin columns unsuitable for large preps. For midis and maxis, the use of gravity flow columns and isopropanol precipitation of plasmid DNA is followed by centrifugal spins and washes and elution as above. In this work, the JetQuick mini- and midiprep kits from Genomed, and the WizardSV plasmid miniprep kit from Promega were used. All maxipreps were made using the EndoFree Maxi kit from Qiagen, giving endotoxin-free preps of high purity. For detailed experimental descriptions, please see http://www.qiagen.com. For most centrifugation steps in the above protocols, the Sorvall RC5C Plus centrifuge was used, with either SS-34 (maximum 30 ml tubes) or HS-5 rotors (maximum 250 ml tubes).

2.1.3 Spectrophotometric quantitation of nucleic acids

A spectrophotometer reads the intensity of both incident and transmitted monochromatic light of a sample compared to a blank control, enabling quantitation of the content when the sample's extinction coefficient, the size of the container/cuvette and sample dilution is known. (According to the *Lambert-Beer law*, relating path length, incident and transmitted light in a closed sample volume.) Using a fixed path length, the optical density or absorbance A is directly proportional to the concentration of the absorbing solute. DNA and RNA have absorbance maxima at 260 and 280 nm, respectively, caused by the purine and pyrimidine rings' ability to absorb UV light. (At 260 nm, an absorbance value of 1 equals 50 μg/ml dsDNA and 40 μg/ml RNA and ssDNA in the sample, respectively). Contaminating protein in a nucleic acid sample has a maximum at 280 nm, (caused by the tryptophane and threonine side chains in the proteins), making the A_{260}/A_{280} ratio a measure of DNA/RNA sample purity. Pure DNA/RNA preparations have a A_{260}/A_{280} ratio of 1.8 and 2.0 respectively, depending upon pH and ionic strength of the liquid used for sample dilution and reference (usually H_2O or Tris-EDTA pH7.5, (TE)). Lower pH results in a lower ratio and reduced sensitivity to protein contamination. It should be mentioned, however, that determination of the RNA concentration in a sample demands a buffer with neutral pH, since the relationship between absorbance and concentration (see above) is based on an *extinction coefficient* calculated at neutral pH (Wilfinger et al., 1997); (Sambrook, 2001). In this work, the Pharmacia Gene Quant apparatus with quartz cuvettes was used for all measurements. 2 μl of the sample was diluted 1:35 or more in 1x TE, pH 7.5, or sterile water.

2.1.4 Restriction enzyme analysis of DNA

Class II-restriction enzymes recognize *palindromic* sequences in dsDNA, and exert their function as an enzymatic defence towards viral DNA in prokaryotes, from where they have been isolated in general. The bacterial DNA itself is well protected towards these endonucleases (cuts within), by DNA-methyltransferases that protect the cell's own DNA from cleavage by the restriction enzyme. Methyltransferases recognize the same DNA sequence as the restriction enzyme that they accompany, but instead of cleaving the sequence, they *methylate* one of the bases in each of the DNA strands (often adenine (A) or cytosine (C)). The methyl groups protrude into the major groove of DNA at the binding site and

prevent the restriction enzyme from acting upon it. Of the three restriction enzyme classes known, number II is the most important and most frequently used in molecular cloning. In addition to the endonucleases, an enzyme class named exonucleases (cuts from ends) is of much use. Some class II-endonucleases leave blunt ends (see Figure 2.1), whilst the majority of the commercially available enzymes leave protruding, staggered ends behind. It is well known that enzyme-catalyzed ligation of single stranded complementary overhangs, or "cohesive ends" is easier than for the blunt. As a standard measure, 1 unit (U) of a given enzyme is per definition the necessary amount of enzyme needed to cut 1 µg of phage λ DNA in a 1 hour reaction at the ideal temperature for the actual enzyme. Every procedure needs optimization, with adjustments of carrier protein and salts and temperature as the most important aspects. Performing multiple restrictions is possible, depending on internal distance between sites of digestion and buffer preferences for both enzymes in addition to the above.

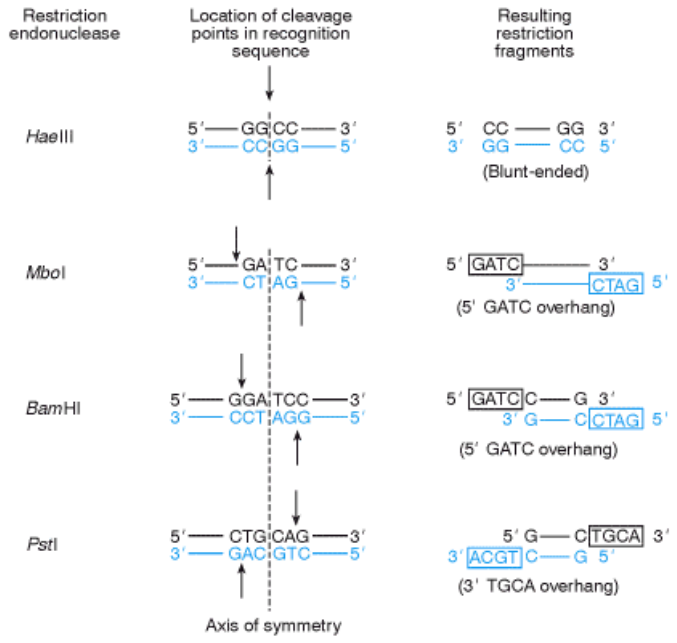


Figure 2.1: Restriction endonucleases can generate blunt-ended fragments or "sticky" fragments with 5′- or 3′-overhanging ends (Read, 1999).

2.1.5 Restriction analysis of plasmid DNA

Standard protocol

A typical reaction mix would consist of restriction buffer, carrier protein, (most often BSA to a final 0,1 mg/ml, necessary for some enzymes), 0,1-5 μg of DNA, 3 Units of enzyme(s) per μg of DNA and distilled or sterile H_2O to the total volume. The DNA fragments can be double stranded PCR products, cloned cDNA or vector DNA. Regarding incubation time and temperature, all enzymes were used according to the supplier's descriptions. Please see the detailed illustration for each cloning procedure further on.

2.1.6 Ligation of DNA fragments using T4 DNA ligase

The bacterial virus T4 harbours a DNA ligase, an enzyme which links juxtaposed 3'-hydroxyl and 5'-phosphate termini of fragmented nucleic acids to form one continous strand, in addition to repairing single stranded nicks in duplex DNA, RNA and RNA/DNA hybrids (Engler, 1982). In this work, T4 DNA ligase was used to ligate pre-digested DNA fragments, i.e. plasmids and PCR products pre-cut with the same enzymes. After incubation, agarose gel analysis of both sample and negative control (plasmid without added ligase), was performed.

Standard protocol

Mix T4 ligase buffer[1], 2-400 Units T4 DNA ligase, 50-300 ng insert DNA and 100-300 ng plasmid DNA plus distilled or sterile H_2O to 15μl. The blend is gently mixed in an eppendorf tube and incubated over night (14-18 hours), at 14-16°C or 1-2 hours at room temperature. If room temperature is preferred for convenience reasons, this may negatively influence the reaction's rate of accuracy, although the T4 ligase enzyme is relatively tolerant when buffer composition is otherwise correct. The recommended DNA concentration for ligation of sticky ends is 0,1-1μM (5'termini). For blunt-end ligation, over night incubation is recommended. When all buffer conditions are correct, the DNA purity and concentration, more so than volume is important. Heat inactivation of the ligase is often recommended, but was not used in this work since the recombinant vectors always were transformed immediately after

ligation. Typically 2-3 µl of the ligase mix was mixed with loading buffer and loaded on a 1 % agarose gel for verification of expected migration lengths. After confirmation, transformation was performed according to section 2.1.1, using 5-10 µl of ligation mix.

2.2 Electrophoresis

Electrophoresis is movement of charged biomolecules in an applied external electric field. There are several suitable media, of which gels are the most commonly used for nucleic acids. Four types of gels have been used in this work, of which two have been used for RNA- and DNA-fragment-separation.

2.2.1 Agarose gel electrophoresis

Agarose is a linear polysaccharide made up of the basic repeat unit agarobiose, which comprises alternating units of galactose and 3,6-anhydrogalactose. It forms gels with pores ranging from 100-300 nm in diameter; ideally suited for low-resolution separation of relatively large molecules. When mixed with an appropriate buffer and heated until the powder has dissolved, the pore size mirrors the agarose content, which varies from 0,2-5% in general, with 1% as the most common, separating fragments from 300 bp-12 kb with ease, see table 2.1.

Table 2.1: The range of separation of linear DNA molecules in 0,3-2,0% agarose gels.

Amount of agarose in gel (% [w/v])	Effective range of separation of linear DNA molecules (kb)
0.3	5-60
0.6	1-20
0.7	0.8-10
0.9	0.5-7
1.2	0.4-6
1.5	0.2-3
2.0	0.1-2

Data adapted from (Sambrook, 2001).

Charged biomolecules will migrate through the pores towards the oppositely charged electrode, their rate of migration through the electric field depends largely on the strength of the field, size and shape of the molecules, relative hydrophobicity of the samples, and on the

[1] 1 x T4 DNA ligase-buffer: 50 mM Tris-HCl pH 7,5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 µg/ml BSA

ionic strength and temperature of the buffer in which the molecules are moving. (Nucleic acids have a net negative charge, and will therefore migrate towards the positive electrode.) In all forms of electrophoresis, the force moving the macromolecule (nucleic acids or proteins), is the electrical potential, *E*. The electrophoretic mobility (E.M) of an ion is known as the ratio of the velocity of the particle, *v*, to the electrical potential: $\mathbf{E.M} = \boldsymbol{v / E}$. Electrophoretic mobility is also equal to the net charge of the molecule, *Z*, divided by the frictional coefficient, *f*: $\mathbf{E.M} = \boldsymbol{Z / f}$. The current in the solution between the electrodes is conducted mainly by the buffer ions with a small proportion being conducted by the sample ions. Ohm's law expresses the relationship between current (*I*), voltage (*V*), and resistance (*R*): $\boldsymbol{R = V / I}$. This equation demonstrates that it is possible to accelerate an electrophoretic separation by increasing the applied voltage V, which would result in a corresponding increase in the current flow. The distance migrated will be proportional to both current and time. However, increasing the voltage would ignore one of the major problems for most forms of electrophoresis, namely the generation of heat. During electrophoresis the power (*W*, watts) generated in the supporting medium is given by $\boldsymbol{W = R \times I^2}$. Most of this power generated is dissipated as heat. Heating of the electrophoretic medium can have the following effects: An broadening of the separated samples; mixing of separated samples caused by formation of convection currents, and thermal instability of heat-sensitive samples which may include denaturation of proteins or loss of activity of enzymes. Finally, a decrease of buffer viscosity and hence a reduction in the resistance of the medium may occur. The evident conclusion must therefore be to follow the golden middle path of low voltage and physiological pH for optimal agarose electrophoresis-assisted separation of nucleic acids. There are, of course, exceptions from the rule.

The *intercalating agent* Ethidium bromide, EtBr, was used for visualisation purposes. EtBr binds between DNA base pairs and fluoresces when activated by UV radiation. To ease the sample loading procedure, the samples are mixed with an appropriate dye-containing loading buffer. The dye serves two purposes; it visualizes of the migration front and facilitates sample loading. (The buffer dye migrates faster than all samples, creating a colour front in the transparent gel, making it possible to estimate when the electrophoresis should be stopped). The loading buffer contains glycerol, making the blend heavier and thus easier to load into the wells of a gel that has been immersed in running buffer. Sample size (base pairs) is assessed using standards. In this work, the phage lambda DNA was cut with the HindIII or BstEII restriction enzymes, to give two useful broad range size standards. We also used two commercial standards; the 1kb and the 100 bp ladders from New England Biolabs.

Protocol (1 % TAE- gel)

1 g of 99,9% pure agarose powder is boiled in 100 ml of 1x Tris-acetate-Ethylene Diamine Tetra-acetic Acid (EDTA) buffer until the powder is completely dissolved, typically 2-3 minutes. The liquid is thereafter cooled until it holds ≈ 60°C, then EtBr may be added to 0,1-0,2 μg/ml, and the liquid poured into a plastic casting tray with pre-inserted comb. The comb once removed, samples are mixed with an appropriate loading buffer and loaded into the wells of a gel that previously has been immersed in running buffer. Gels are run in 1x TAE buffer at 60-100 V until the coloured dye has migrated three quarters of the gel, and then photographed using UV illumination. In this work, the bands were photographed using the BIO RAD GelDoc 1000.

2.2.2 Polyacrylamide gel electrophoresis

Raymond and Weintraub introduced the polyacrylamide gel electrophoresis (PAGE) technique in 1959. This high-resolution electrophoresis method is able to separate single-stranded DNA molecules that differ in length by just one nucleotide, making it amenable to nucleic acid sequence determination as well as separation of nucleic acids and proteins. A polyacrylamide gel consists of chains of acrylamide monomers, cross-linked with N'N'methylenebisacrylamide, or "bis" units. Pore size depends both upon the total concentration of monomers, (including bis), and the ratio of acrylamide to bis, typically 19:1 for a sequencing gel. The separation of proteins and peptides takes place in gels typically ranging from 8-15% PA content, whilst nucleic acids are electrophoresed in gels with 5-12,5% [w/v] acrylamide, see table 2.2. The cross-linking reaction is oxygen-sensitive and polymerisation needs Ammonium Persulfate (APS), and TEMED (N',N',N',N-tetramethylethylenediamine), as initiator and catalyst, respectively. When nucleic acids are separated in PA gels, different MW-sizes are used due to the narrower resolution range of the PA-gels. The φX174 DNA cut with HindIII was used as MW-standard upon electrophoresis of nucleic acids in this work. (Protein gel electrophoresis is discussed in more detail in section 2.2.3.) Various additives serve alternative purposes, as the addition of urea denaturates duplex DNA in a sequencing gel, and the addition of the detergent Sodium Dodecyl Sulphate, (SDS), in protein PA-gels where we wish to focus on protein size; not charge. (SDS binds to and

unfolds the protein, creating an overall negative charge independent of native dimensions, but largely proportional to peptide length.) For protein separation, we often wish to optimize the resolution of the separating gel. This can be achieved with a glycine/Cl^-stacking gel, making all samples enter the separating gel at the same time point. When the electric field is established, chloride, protein, bromophenol and glycinate ions all begin to move toward the anode. Due to the low pH, glycinate ions (weak acids with pK_a below the pH of the stacking gel buffer) will be present as their equivalent non-ionic zwitterions that are immobile. Failure of glycine zwitterions to move into the stacking gel creates a deficiency of mobile ions. However, a constant current *must* be maintained throughout the entire electrical system. This is accomplished in the area between the leading chloride ions and the trailing glycinate ions by an increase in voltage. The result is a very high voltage gradient occurring between the chloride ions and the glycinate ions. In this condition the relative ion mobilities are glycinate<protein<bromophenol blue<Cl^-. In this strong local electric field the anionic proteins all migrate rapidly through the large pores of the stacking gel which do not impede their progress. If any of the proteins overtake the chloride ions, they slow down because wherever there are chloride ions there is no ion deficiency. Thus all samples enter the separating gel at the same time, and the higher pH reforms glycinate ions. This balances the ion deficiency. Additional details and protocols for SDS-PAGE on protein samples can be found in section 2.2.3.

Table 2.2: The range of separation of linear DNA molecules in 3,5-20,0 %[w/v] acrylamide gels.

Acrylamide (% [w/v])[a]	Effective range of separation (bp)	Xylene cyanol FF[b]	Bromophenol blue[b]
3.5	1000-2000	460	100
5.0	80-500	260	65
8.0	60-400	160	45
12.0	40-200	70	20
15.0	25-150	60	15
20.0	6-100	45	12

[a] N,N'-methylenebisacrylamide is included at 1/30th the concentration of acrylamide.

[b] The numbers given are the approximate sizes (in nucleotide pairs) of fragments of double-stranded DNA with which the dye co-migrates.

Data adapted from (Sambrook, 2001).

Standard Protocol for 7,5% Polyacrylamide gel

Work in a ventilated hood, and always clean the glass plates with isopropanol before assembly. Mix the reagents from table 2.3 adding TEMED as the final ingredient.

Table 2.3: Standard recipe for a 7,5% PA gel.

Acrylamide 40%	1,125 ml
1 x Tris-Acetate-EDTA (TAE)	4,8 ml
APS 20%	45 µl
TEMED	10 µl

Swirl container to mix contents and pour into the premade plastic or glassware form. Finish by inserting a clean comb of choice. The gels were run at 200 V at room temperature in 1 x TAE buffer in electrophoresis equipment from BIO-RAD. DNA and RNA-fragments were visualized using 3-6 µl of the *intercalating dye* SYBR Gold (Molecular Probes). This dye is comparable to Ethidium bromide, intercalating between base pairs in the nucleotide sequence. It is however less toxic and more sensitive to light. It is most often added to 5-10 ml running buffer containing the gel after the run, in contrast to EtBr.

2.2.3 Electrophoresis of proteins

The percentage of acrylamide in a PA gel determines its resolution range (table 2.4).

Table 2.4: Relation between acrylamide percentage and range of protein separation for a PA gel.

Acrylamide percentage	**Range of separation**
8%	36-94 kDa (7,5%)
12%	15-50 kDa
15%	12-43 kDa

Data adapted from (Sambrook, 2001).

We chose 12 and 15% gels since our expression products had molecular weights of 21-28 kDa. The gels were cast as described for a 7,5 % PA gel, using the recipe from table 2.5.

Table 2.5: Recipe for 12% SDS PA gel with 5% concentrating gel.

12% separating SDS PA-gel:

Acrylamide 30%	2,0 ml
1M Tris pH 8,8	1,9 ml
SDS 10%	50 µl
APS 20%	25 µl
H_2O	1,1 ml
TEMED	2 µl

5% concentrating gel:

Acrylamide 30%	0,5 ml
1 Mtris pH 6,8	0,67 ml
SDS 10%	40 µl
APS 20%	20 µl
H_2O	2,7 ml
TEMED	4 µl

10 µl of supernatant was mixed with an equal amount of 2x loading buffer, while pellets were dissolved in 20 µl loading buffer. All samples were boiled at 95°C for 5 minutes, spun briefly in a benchtop centrifuge and loaded onto a 12% SDS-PA gel with slim pipette tips. Protein molecular weight standards are often pre-stained for ease of detection. In this work, two such commercially available standards were used; the broad range (6,7-208 kDa), and the dual color 'Precision Plus' (10-250 kDa), both from BIO-RAD. The gels were run at 15mA for 30 minutes, then 30 mA to end at room temperature. Proteins were fixed for 15 minutes in a Coomassie fixing solution, stained with Coomassie Brilliant Blue staining solution for 60 minutes, and destained over night using the fixing solution. (Please see the appendix for all recipes.) Coomassie Brilliant Blue has a downward resolution of 0,1 µg or 2 pmol protein for visualization in a gel, making this a crude visualisation method applicable to bacterial lysates.

2.2.4 Purification of DNA from agarose gel slices

Isolation of DNA fragments from agarose gel slices was performed using the QiaQuick Gel Extraction Kit from Qiagen. The slice of gel containing the band of interest is resuspended in a saline buffer at 50-60°C. The buffer in question contains a pH-sensitive dye in addition to guanidine thiocyanate (a strong protein denaturant) and *chaotropic* salts. The solution is loaded onto a spin column with a DNA-binding silica-membrane, which eases further EtOH-wash and elution using 10mM Tris-Cl or Tris-EDTA (TE) pH 7,5-8,5. For a detailed protocol, please see the Qiaquick handbook at http://www.qiagen.com.

2.3 Polymerase Chain Reaction

The PCR method enables amplification of nucleic acid sequence fragments in vitro by primer extension (Mullis et al., 1986). Briefly, a pure DNA fragment or equivalent material from other non-purified sources such as bacterial colonies or phage plaques can be amplified *in vitro*, by mimicking the *in vivo* DNA polymerase replication machinery. The cyclic process includes three repeated temperature-regulated levels plus pre- and post-preparation steps. First, the samples containing the double strand fragment of interest is ***denatured*** at 95°C or higher for 15-240 seconds, permitting 18-24 bp sequence-specific oligomers called primers to attach to the single stranded template at the lower ***annealing*** temperature at step two, lasting for 30-60 seconds. (The latter temperature can be estimated by various methods, but often equals the template-primer melting temperature (T_m) -5 degrees, ranging from 40 to 60°C, and needs to be optimized in most cases.) An included DNA polymerase attaches at the primer site where DNA is double stranded, and completes the remaining sequence in the ***elongation*** step. This temperature is the optimal for the chosen enzyme, and usually lies around 72°C. Deoxynucleotide trisphosphates, specific salts needed for enzymatic activity ($MgCl_2$ or $MgSO_4$), and necessary additives ensure proper PCR product formation. The cycle is repeated typically 15-40 times in a thermal cycler. Since amplified products from the previous cycle serve as templates for the next cycle of amplification, PCR is an exponential process, following an exponential curve from cycle #2. At each step, 2^n new copies of the template are produced, where n is the number of cycles. PCR can be illustrated with the formula $N= N_0 \times 2^n$, where N is the number of amplified molecules; N_0 represents the original molecule number and n the number of cycles. In a typical reaction, 10^4-10^6 copies of the initial DNA

template will be produced. PCR is a sensitive and potent tool used in all sorts of biological research, and has been utilised extensively in the DNA work of this study. A number of factors can improve the reaction, typically primer- and template design and template amount as well as magnesium concentration. Adding formamide, betaine, DMSO and/or glycerol have been successful. (*Their proposed mechanism is to lower the DNA melting temperature, thereby enabling improved elongation and accuracy. Fully melted DNA hinders elongational arrest or enzyme throw-off caused by difficult secondary structures.)* The magnesium concentration directly affects the polymerase activity, making metal complexing salt buffers with high EDTA content unwanted as DNA dilution liquid. The $MgCl_2$ or $MgSO_4$ concentrations should be optimized for each individual reaction. Generally, excess Mg^{2+} will result in the accumulation of non-specific amplification products and insufficient Mg^{2+} will reduce the yield of wanted PCR product. Adding Bovine Serum Albumine (BSA) and non-ionic detergents such as Tween20 may stabilise the enzyme. The identity of the latter is often critical, and temperature-activated, or so-called thermostable polymerases, are preferred. Some, like the PfU polymerases (derived from the thermophilic archae *Pyrococcus furiosus*), have 3'→ 5' proofreading activity in addition to the 5'→ 3' activity provided by the *Taq* polymerases, derived from the thermophilic bacterium *Thermus aquaticus*. Most commercial heat-stable polymerases may be combined with an antibody that binds to the enzyme at lower temperatures, thus inhibiting the polymerase activity until the tubes have been put into the thermal cycler. This *hot start*-method is the most common, minimizing the risks and amount of labour needed. Throughout this work, both heat-stable *Taq* and *PfU* polymerases have been provided by Dr. David J. Warren, The Norwegian Radium Hospital.

A standard PCR reaction mix:
(All concentrations are final. $MgCl_2$, primers and buffer were vortexed prior to use.)

1-50 ng of template DNA
1 x PCR buffer (10 mM Tris-HCl, pH 8.8, 50mM KCl (or NaCl), 0,08% (v/v) NP-40)
2,5 mM $MgCl_2$
0,1-0,5 μM of each primer
0,2mM dNTP
1-2 Units of thermostable polymerase

Volumes usually range from 20-100 μl. If the $MgCl_2$ stock solution has been through many freeze/thaw-cycles, it may be heated to 90°C for 10 minutes to restore the homogeneity of the solution, (Hu et al., 1992). All PCRs in this work were done using the GeneAmp^R PCR System 2400/9700 with 0,2 ml tubes (both from Perkin Elmer).

2.3.1 Handling of PCR products

PCR reactions should be set up in a designated area, preferably an isolated room where the products will be stored later. Volatile PCR products (aerosols) can contaminate reactions and laboratory equipment, and sample tubes should therefore be kept closed or opened only in "secure" areas (hoods, clean benchtops etc), and separate pipettes should be used for products.

2.3.2 Primer design

The choice of primers for plasmid constructs depends on the identity of the template. Precautions ought to be taken to avoid primer-dimer formation and hence poor performance, for instance a high GC-content at the 5'end, and neither non-complementary sequences nor high GC content at the 3'-ends. Restriction enzyme-sites can be added to the template using 5' *overhanging primer ends*, which will be used for later cloning purposes. It is common to use overhangs with such sites and subsequently cut the PCR product with appropriate restriction enzymes before inserting the readily cut PCR product into a vector of choice. This method was used several times in this work. Estimation of the melting temperature T_m is crucial, and various computational methods exist. (At T_m, 50% of the primer and its complementary single strand sequence are present in a DNA duplex.) Most important; cycling conditions and buffer concentrations should be optimized for each primer pair.

2.3.3 Purification and analysis of PCR products

PCR products were purified using DNA-binding spin column-based kits from Qiagen. A guanidine hydrochloride/isopropanol-buffer is mixed with the PCR reaction, and the blend loaded onto a spin column. Basically, denaturated impurities pass through the DNA-binding silica membrane, whereas the amplified DNA does not. EtOH-wash and elution in a 10 mM Tris or TE pH 8,5-solution follows to give a reasonably pure preparation. This method is not as thorough as purification from gel (because other DNA fragments are not removed during the spin column purification), and should be avoided when further high purity-applications (sequencing, blunt end cloning) are considered. Amplified PCR products were analysed using 1,5-2% agarose gels or polyacrylamide electrophoresis (PAGE) as described earlier. The products were always stored at 4°C.

2.3.4 Colony PCR for transcription/translation analysis

A quick and easy way of screening a large number of bacterial colonies is by direct PCR, evading the purification prep and further agarose gel electrophoresis of a good number of possibly negative samples. The template for each reaction is simply a minuscule amount of each colony to be tested. A pipette tip or a sterile toothpick may be used to touch a colony of choice, and the tip is left in sH_2O for the template to dissolve. The tip is removed, and the template-containing water mixed with remaining PCR reaction ingredients. Insert-specific primers should be used in a standard PCR as described in previous sections. For optimal results, it is common to denaturate the bacterial template at 94-99°C for 4 minutes before the premier PCR cycle. An important point is the amount of template loaded, as too much bacteria will easily inhibit the PCR. 0,7-2% agarose- or PAGE electrophoresis can be applied to 2-5 µl of the PCR reaction.

2.3.5 End phosphorylation of a PCR product with T4 polynucleotide kinase

The polynucleotide kinase (PNK) from the bacteriophage T4 catalyzes the transfer and exchange of P_i (inorganic phosphate), from the γ position of ATP to the 5´ hydroxyl terminus of polynucleotides (duplex and single-stranded DNA and RNA), and nucleoside 3´-monophosphates. PNK also catalyzes the removal of 3´- phosphoryl groups from 3´-phosphoryl polynucleotides, deoxynucleoside 3´-monophosphates and deoxynucleoside 3´-diphosphates (Engler, 1982). T4 PNK was used for radioactive labelling in this work, please refer to section 2.6.6. For non-radioactive phosphorylation, up to 300 pmol of 5´ termini in a 50 µl reaction containing 1 x T4 PNK Reaction Buffer[2], 1 mM ATP and 10 units of T4 PNK is recommended. The efficiencies of blunt and recessed 5´-end phosphorylation can be improved by heating to 70°C for 5 minutes, then chilling on ice prior to kinase addition and by adding PEG-8 to 5% (w/v) (Sambrook, 2001). Since PNK is inhibited by ammonium ions, DNA was not precipitated in the presence of ammonium ions prior to phosphorylation. The following protocol was used for the creation of the pFastBacHTa/*COAS2* vector, but the method is applicable in general to all kinds of blunt-end PCR products.

[2] (70 mM Tris-HCl, 10mM $MgCl_2$, 5 mM DTT)

Standard protocol

2 µg in a total of 71 µl of H_2O was heated to 70 °C for 5 minutes and then chilled on ice. 10 µl of (700 mM Tris, 100 mM $MgCl_2$) pH 7.6 was added and resuspended before the addition of fresh 5 µl 100 mM DTT and 10 µl 10 mM AdenosineTriPhosphate, (ATP). 4 µl (40 U) of T4 PNK was added. After gentle mixing, the reaction mix was incubated for 30 minutes at 37 °C. The products were purified according to sections 2.3.3 and 2.2.4.

2.4 Sequencing

The dideoxy, or chain termination method from Sanger (Sanger et al., 1977) was used along with automated MEGA BACE 500 sequencing kits and DYEnamic™ ET dye reagents from Amersham to determine nucleotide sequences. In the dideoxy method, the fluorophore-tagged 2'3'dideoxy analogues of the 3'deoxy nucleoside trisphosphates act as specific chain-terminating inhibitors for the included DNA polymerase. Each NTP has an individual tag, (e.g. red fluorochrome for ATP, blue for GTP), improving the assay accuracy and enabling all reactions to take place in a single tube. As mentioned, a commercial sequencing kit was used, but the producer has not published the composition of the DYEnamic mastermix. The reaction is mainly a single 30-cycle PCR where primer and template is mixed with the mentioned mastermix containing ddNTPs, Mg^{2+}, a saline buffer (probably Tris-HCl and alkaline salts), and polymerase. The product is precipitated with 7,5 M ammonium acetate and 70% EtOH, centrifuged, washed and air-dried. The pellet is resuspended in 10 µl loading solution before the run (70% formamide, 1mM EDTA). The MEGA BACE instrument performs capillary electrophoresis, where the sample electrophoretically migrates through tubes with separation matrix, (3% linear polyacrylamide in 1x Tris-Borate-EDTA (TBE)). The laser beam scans the capillaries continuously, exciting the fluorochrome tags. The emitted photons are filtered through six beamsplitters, leading to their respective photomultiplier tubes (PMTs). The light is converted to electrical pulses, that again are digitized to give an electropherogram (raw data), for each capillary. Taken together, the raw data gives us the readthrough sequence for the user to analyze with the Sequence Analyzer software v3.0 from Molecular Dynamics. In a typical plasmid sequencing reaction, vector primers flanking both ends of the gene of interest were used along with 400 ng of template,

generating products of 5-800 bases in each reaction. For a more detailed sequencing PCR procedure,please see http://www5.amershambiosciences.com.

2.5 Sequence analysis

All database searches for either nucleotide- or amino acid sequences including homology searches were performed using variants of the Basic Local Aligment Search Tool (BLAST) at the National Center for Biotechnology Information (NCBI) website (Altschul et al., 1990). Aligning a pair of nucleotide sequences was done using the Blast2 tool, which finds multiple local alignments between the two sequences (Tatusova and Madden, 1999). The primers used in this work were created with the Informax Vector NTI suite 8. Additional protein and mRNA knowledge was extracted with various online services; please refer to the appendix section. All online services used in this work are accessable through URLs listed in the appendix.

2.6 Expression analysis

Working with RNA

Ribonucleic acids are widespread molecules with multiple and dynamic functions. Working with and purifying RNA demands that some precautions should be made. First, all samples and RNA equipment should be shielded from RNAses, the latter existing e.g. on glassware and fingers. Second, the RNA should be kept on ice when possible, and stored at -80°C. Water was treated with diethylpyrocarbonate (DEPC) before autoclaving. DEPC reacts with histidine residues in the enzymes, thereby inhibiting their harmful activity. After autoclaving, DEPC is hydrolyzed to form CO_2 and EtOH. All of the equipment, except glass, was treated with RNAseZAP before use to remove all traces of RNAse. This is a solution from Sigma of unknown exact composition, but its material safety data sheet states the existence of three different agents, of which some have oxidizing effects, with known inhibitory effects on RNAses. Electrophoresis trays were treated with SDS and DEPC-water between runs. Compounds with primary amine groups, e.g. Tris, may not, however, be DEPC-treated as they react with and thus inhibit DEPC. (See further details on RNAses and DEPC-treatment at http://www.ambion.com/techlib/tb/tb_178.html.) Most importantly, preferentially non

powdered gloves were always used and changed often when handling RNA.

2.6.1 Isolation of total RNA

For all applications, the biochemical Trizol from Invitrogen was used to isolate Total RNA from cell material. As described by Chomczynski and Sacchi (1987), the monophasic reagent contains the carcinogenic phenol in mixture with guanidinum thiocyanate, (GTC), a toxic, protein-denaturating compound. Trichloromethane (chloroform) is added to create two phases after agitation. The organic phase holds denaturated proteins and DNA, whilst RNA and some contaminating DNA are present in the aqueous phase. (DNA levels are kept low if the amount of Trizol is increased. Here, 1,2 ml Trizol was used for 10^7 cells). After centrifugation, total RNA is precipitated with isopropanol, followed by washes with 75% ethanol, air-drying of the pellet and resuspension in a proper solvent. At the time of harvest of the cell flask, the cell pellets were most often resuspended in Trizol and the container immediately snap-frozen in liquid N_2. Working in a ventilated hood, the cell pellets in Trizol reagent were thawed on ice and then homogenised with a pipette. Because the RNA would be used for micro array analysis after a short pause in the freezer, sterile water was always used to dissolve the pellet, even though TE or *sarcosyl* may be preferable to limit the extent of degradation and RNA loss. It is important to register that TE may not be garanteed to be free of RNAses (please see previous paragraph). All RNA samples were quantified with the Pharmacia Gene Quant apparatus, or the NanoDrop apparatus (Nano Drop Technologies). The latter has a detection range from 2 ng/µl-3,7 µg/µl, and needs merely 1 µl for quantification. 10 million cells would give 350-400 µg of RNA, typically. For detailed and updated protocols, please refer to the suppliers' homepage at http://www.nanodrop.com. All RNA samples were assayed for quality with the Agilent BioAnalyzer. This advanced miniature gel electrophoresis system also holds quantitation functions. Solutions between 0,05 and 0,4 µg/µl RNA may be used, requiring merely 1 µl for a run in the predyed gel system. The protocol from the supplier for the RNA 6000 Nano Chip™ was followed. Please see http://www.agilent.com for further information regarding the BioAnalyzer and its applications.

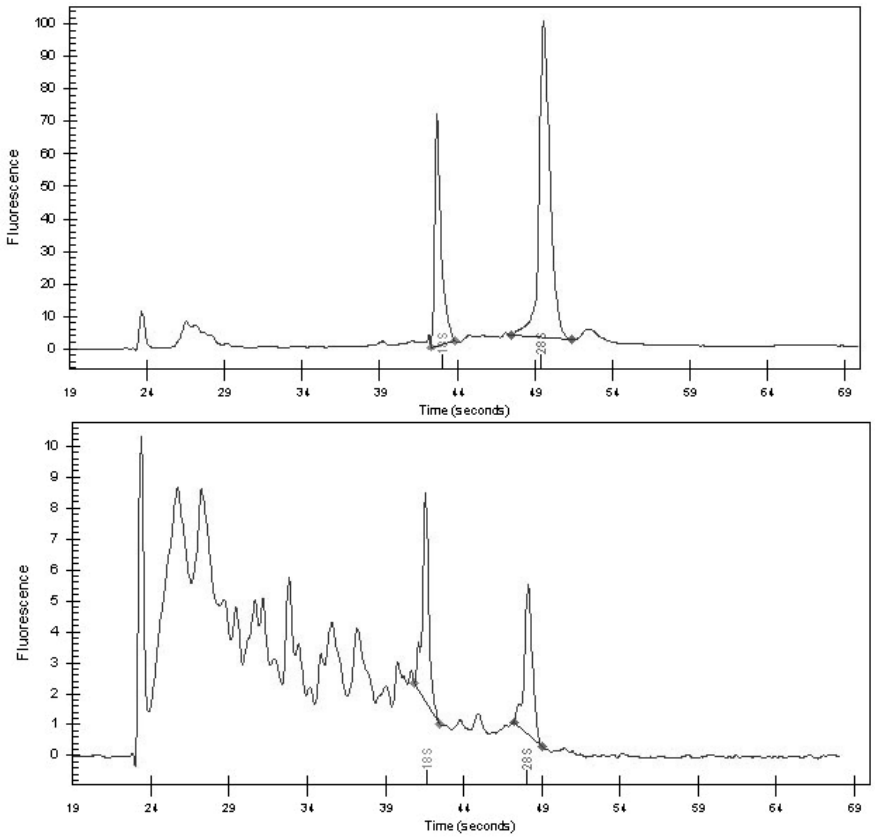


Figure 2.2: Picture from the BioAnalyzer software showing a pure eukaryotic total RNA preparation (upper case) and a degraded sample (lower case). The peaks at 24, 44 and 49 seconds represent 5S, 18S and 28S RNA, respectively.

2.6.2 Introduction to the northern blotting method

Expression analysis at the RNA level is mainly focused on detection of post-transcriptionally modified mRNAs in a sample. In the northern blotting technique, negatively charged, separated RNA preparations are transferred from an agarose gel to a positive nylon membrane by capillary forces (see figure 2.3). After transfer, the nucleic acids can be fixed to the membrane using either brief ultraviolet (UV) irradiation or heating at 80°C for about half an hour. The RNA filter can now be hybridized to a radioactive-labelled RNA/DNA probe in a strong saline solution at a fixed pH and temperature. Autoradiography gives hybridised bands corresponding to the specific mRNA in each sample.

2.6.3 Agarose gel electrophoresis of RNA

Messenger RNA (mRNA) is much smaller than genomic DNA, so it can be analyzed by agarose gel electrophoresis without the enzymatic digestion steps that are necessary for the analysis of high molecular weight DNA applied in the Southern blot method. 1% agarose gels containing the denaturant formaldehyde are used to separate RNA electrophoretically. Formaldehyde forms unstable Schiff bases with the single imino group of guanine residues, thereby preventing intrastrand base pairing. The binding of RNA to formaldehyde is strongly affected by changes in pH (Sambrook, 2001). During electrophoresis, a pH-gradient is established in the buffer system, making formaldehyde-assisted denaturation of RNA difficult. This negative effect may be prevented by recirculation of the buffer through the formaldehyde gel during electrophoresis. For a 300 ml gel, 3 g agarose, 6 ml 1M sodium dihydrogenphospate (NaH_2PO_4) pH 6.6, and 240 ml DEPC-treated H_2O are mixed and heated in a microwave oven to dissolve the agarose. After cooling to ≈ 60°C, 54 ml formaldehyde and 4,5 µl of 10 mg/ml EtBr are added in a secure ventilated hood to avoid the inhalation of harmful vapours. In general 2-10 µg of total RNA is loaded onto each lane of the agarose gel (after denaturation at 65-70°C for 10 minutes), using a loading buffer containing formamide. The running buffer (20 mM NaH_2PO_4 pH6.6) is recirculated in the formaldehyde gel during the 10 V run, thus maintaining a stable pH throughout the gel over night. The gel is run at 30-35 V until the samples have entered the gel, and then at 10 V until the next morning. At this point, the electric pump was switched on for buffer recirculation and the jet properly adjusted.

In this work, the gel was visualized with the GelDoc instrument (BIO-RAD) after electrophoresis, followed by a 30 minute-wash in 10 x Standard Saline Citrate (SSC).

2.6.4 Blotting from gel to filter

Transfer of RNA from the gel and onto the nylon membrane over night was performed as illustrated, placing the gel upside down on a buffer-soaked pad of Whatman filter paper. On top of the gel, we put a Hybond N^+ nylon filter (Amersham Pharmacia Biotech), and another Whatman paper again followed by large amounts of dry absorbing paper with a small weight to ensure tight transfer. The RNA is carried along through the gel and onto the nylon filter as the mobile phase, (10 x SSC buffer), migrates upwards, towards the weight, see illustration below.

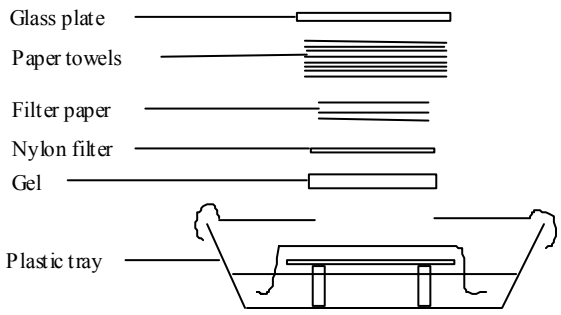


Figure 2.3: Schematic representation of a capillary blotting apparatus. Capillary forces mediate transfer of the nucleic acids from a formaldehyde-agarose gel to a filter paper. The mobile phase (saline solution) is drawn towards a weight placed on top of a stack of dry paper towels. Passing slowly through the gel, the solution solubilizes the nucleic acids and transfers the latter to the above filter paper.

2.6.5 Preparation of RNA membranes for hybridisation

After gel visualisation, the wells' positions were indicated on the filter using a soft pencil. (The picture ideally should indicate that no RNA is left in the gel, but this is made difficult because of both the 30' SSC wash and the background caused by remaining formaldehyde.) Fixation of the RNA to the nylon filter was performed using 70 mJ UV light for 45 seconds. The filter was then soaked in 50 mM NaH_2PO_4 pH 7.2, and pre-hybridised in a glass container containing Church hybridisation buffer[3] at the given temperature. The filter is "polar", having most of its RNA on one side. This has to be facing away from the glass to obtain optimal hybridisation results. The container is kept in a rolling incubator at a set temperature.

2.6.6 Probe synthesis and purification

cDNA probes intended for use with the Church hybridisation method should ideally have a final undiluted concentration of 25-50 ng/µl. In this work, 25-50 ng was used for all hybridisations where [α-^{32}P]-dCTP was the radioactive reactant. Accession numbers for all the probes are located in the appendix section.

cDNA clones for the *Homo sapiens* sex determining region Y (SRY)-box transcription factors *L-SOX5, SOX6* and *SOX9* were ordered from the Ressourcenzentrum für Genomforschung

[3] (0,5M Na_2HPO_4 pH7.2, 7% SDS, 1mM EDTA)

(RZPD), Germany, and shipped in LB-agar containing bacteria which harboured the recombinant plasmids. (Unigene accession numbers are located in the appendix section.) Some of the shipped LB-agar was diluted in warm LB-medium, grown over night and 100 µl dispensed on LB-plates containing 50 µg/µl ampicillin. (SOX5 and -6 were present in the pOTB7 and pCMVsport6 plasmids, respectively, that both harbour an ampicillin-resistance gene.) The SOX9 gene was shipped as a cDNA clone in the T7T3PacI plasmid, necessitating 25 µg/µl chloramphenicol for selective growth in LB-medium. All three SOX-factors each gave rise to several clones that were picked on the following day with a sterile pipette tip, dissolved in LB-medium and grown over night to create culture stocks for miniprepping and restriction enzyme analysis. Subsequently, maxipreps were made as described in section 2.1.2. Sequencing confirmed the identities of *SOX5*, *-6* and *-9* at both mini- and maxiprep stages of purification. Probes were made for northern blotting using restriction enzyme digestion of 6 µg plasmid DNA with two appropriate enzymes in each case (35-40 units of each enzyme). 2 hours of incubation at 37°C was promptly followed by gel purification according to section 2.2.4. After quantification, the probes were diluted to a duplex DNA content of 25-50 ng/µl.

A purified COAS2 495 bp cDNA of 32 ng/µl was used for all blots. Another probe used for hybridisation held 1000 bp of the collagen type two alpha one (COL2A1), of 25 ng/ µl.

2.6.7 Probe labelling (Modified from Feinberg and Vogelstein, 1983)

25-50 ng DNA or cDNA fragment was mixed with 1 µl (25 µg/µl) dN_6 primer and sterile water to 10 µl. The blend was boiled for 4 minutes to denature the fragment, and cooled on ice, thereby annealing the primer to the single stranded fragments. Subsequently, 4,5 µl 5 x OLB buffer[4], 1 µl 3mM dNTPs-C, 3 µl "Redivue" [α-^{32}P]-dCTP 3000Ci/mmol solution (Amersham Biosciences, containing a red dye and β-mercapthoethanol), and finally 1 µl of (5U/µl) Klenow polymerase were added before incubation took place at either room temperature for 2-5 hours or at 37°C for 0,5-2 hours. A cleanup after the probe synthesis was performed with Sephadex G50 gel separation columns. 5 µl of NT stop solution (600 µl 25 mg/ml blue dextran, 50 µl 10% SDS, 50 µl 0,5M EDTA) was added before the probe was applied onto the homemade column. The blue fraction, containing the dextran and probe, will

[4] 250 mM Tris-HCl pH 7.6, 25 mM $MgCl_2$, 1 M HEPES

be excluded from the gel pores due to the large MW, and will therefore migrate fast through the column, in contrast to the small molecules in the red fraction. Elution was halted after collection of the blue fraction, followed by a Geiger counter control of the probe. Good incorporation should give 2 times the signal when the column containing non-incorporated [α-^{32}P]-dCTP and the collected probe are measured from the same distance with the Geiger counter, (e.g. 100 cps for the probe, and 50 for the red column, containing non-incorporated [α-^{32}P]-dCTP). After pre-incubation (10-15 minutes), the Church solution (0,5M Na_2HPO_4 pH7.2, 7% SDS, 1mM EDTA) was changed. The probe was heat-denaturated for 4 minutes in boiling water, and immediately added to the warm Church solution. The main hybridisation period typically lasted 12-20 hours.

2.6.8 Wash

The following morning, the buffer was changed for pre-heated Church wash solution, (1%SDS, 40mM Na_2HPO_4 pH7.2). Two 10 minute-washes at 65°C followed by a final 15' wash at RT with gentle shaking were used in general. This could however vary from probe to probe. The remaining radiation was measured with a Geiger counter, and the filter prepared for autoradiography.

2.6.9 18 S rRNA hybridisation

To assay RNA sample loading and enable quantification of the band intensities, an 18S rRNA oligonucleotide was hybridised to the chosen filter, as performed by (Maelandsmo et al., 1996)). All hybridisations and washes were performed exactly as the above, but at 55°C, with shorter incubation times (2-4 hours). A synthetic single strand-oligonucleotide without a 5'PO_4 group was phosphorylated using [γ-^{32}P]-dATP. Labelling with the polynucleotide kinase (PNK) from the bacterial virus T4 takes place at 37°C, and no clean up or heat inactivation of the enzyme is needed. 1 µl (50 ng/µl) 18 S oligo, 1 µl 10 x PNK buffer (recipe found in the appendix), 2-3 µl "Redivue" [γ-^{32}P]-dATP (3000 Ci/mmol, Amersham Biosciences), 1 µl (10 U/µl) T4 PNK and sH_2O to 10 µl was gently mixed and incubated at 37°C for 30-45 minutes. 90 µl of 1 x TE was added to halt the reaction before the incorporation of [γ-^{32}P]-dATP was measured using two Diethylaminoethyl (DEAE) cellulose membranes (NA 45, Schleicher & Schuell). One µl was applied onto each DEAE-filter. Only

filter #1 was washed three times in 5 ml 0,5 M Na_2HPO_4 pH7.2, before both filters were measured using a Geiger counter. The difference between these two will ideally give a ratio of 70-80 % under optimal conditions. Unincorporated nucleotides do not have to be removed prior to incubation. The probe was added to the filter-containing glass tube, and hybridised to the filter as described in this paragraph.

2.6.10 Strip of radioactivelly labelled filters

Removal of radioactive probes before new hybridisations was done regularly in this work. The filter was washed in a rocker for 10 minutes in 80-90°C RNA strip solution (0,1 x SSC, 0,1% SDS), before measurement of radiation and further use or storage at -20°C.

2.6.11 Autoradiography and quantification

The membranes were wrapped in plastic and put with a sheet of sensitive film (X-OMAT UV-film, KODAK), into film cassettes with intensifying screens. After exposure at -80°C, the films were developed in the automated 3M Xp515 developer and bands quantified using the 18S signals. In this work, the Molecular Dynamics Personal Densitometer SI connected to a PC with the GeneQuant™ software was used for the quantification. The northern blot procedure has its weaknesses; only moderately abundant mRNAs can be detected. Total RNA was used for all downstream applications in this work, even though verification of candidate genes after microarray analysis using the northern blot method would benefit from increased sensitivity. To obtain optimal sensitivity, the RNA preparation can be enriched for mRNA, a fraction that ordinarily makes up less than 10% of the total RNA content of a cell or tissue. By removing all RNA molecules that lack the 3´polyadenine tail, the sensitivity will be highly increased (Aviv and Leder, 1972).

2.7 Microarray analysis

There are many types of arrays, from tissue and protein microarrays via carbohydrate and genomic to cDNA and oligonucleotide arrays. Only two-channel cDNA spotted microarrays were used in this work. A cDNA-microarray slide has several thousand spots (Ø < 300 μm) printed in a grid on a coated surface. Each spot contains PCR products or synthetically generated nucleotide sequences, double or single stranded. The cDNA to be probed is made from mRNA in a reverse transcriptase reaction. It consists of test and reference total RNA (e.g. cell line or tumour material). Visualisation of the results is made possible through incorporation of fluorescent dyes (fluorochromes with known excitation and emission spectra), typically Cy3 and Cy5 (in this work from Amersham Biosciences containing NHS-ester leaving groups). When the single strand Cy-dye tagged nucleotide sequence is hybridised to the array, the fluorochrome-tagged cDNA will bind in a stoichiometric fashion to the cDNA in each spot. After wash and scan of the slide, the picture gives a ratio in each spot between the Cy5-and Cy3 channels. This ratio indicates the relative abundance of specific mRNAs in each sample. A Cy3-tagged sample will emit in the green spectrum, whilst Cy5-tagged material will seem red on the computer screen. (The colours have been artificially created, as each raw picture has no colour, i.e. is grey.) A raw image is seldom completely trustworthy due to spot variation, uneven incorporation of dyes and perhaps different cDNA masses in reference and sample. One must therefore filter bad spots and normalize the ratios in order to quantify and analyze the sample's mRNA profile correctly. It is common to apply $\log_2$-transformed ratios, because this creates a model of up- and down-regulated mRNAs in the sample versus the reference that is more easily interpreted. To illustrate an array image, we could imagine a sample where the tumour suppressor gene coding for the p53 protein was deleted, compared to a p53-containing reference. If the sample has been tagged with Cy5-dye, then the lack of p53 mRNA will cause a green spot to occur, because there is only p53 in the Cy3-tagged reference. This example fully illustrates the necessity of p53 (or any other gene of interest for that matter) to be printed on the actual array, and to be expressed in the reference to give a reliable measurement. Otherwise, we would never detect a lack or excess of our gene of interest.

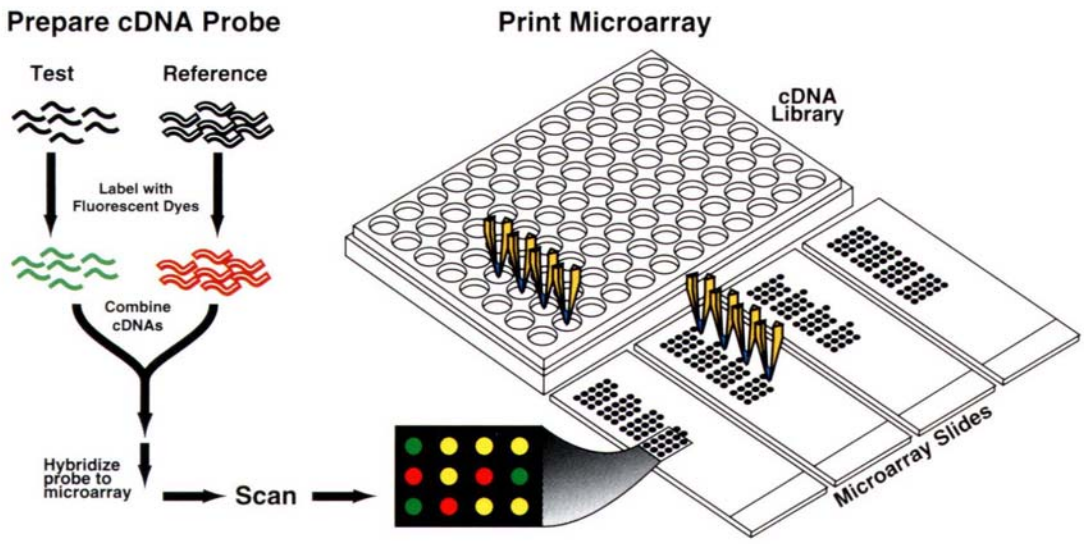


Figure 2.4: The process of microarray hybridisation. The cDNAs are printed in a regular array onto glass slides. The test and reference RNA samples are labeled with different fluorochromes and combined before they are hybridised to the processed microarray. After stringent washes, the microarray is scanned in a laser-scanner device, and the image processed to generate numerical data. Figure from (Meltzer, 2001a).

2.7.1 Probe preparation, hybridisation and wash

Twenty micrograms of test- and reference total RNA was labelled with the indirect aminoallyl labelling technique, which couples the Cy-dyes to the generated cDNA in a two-step procedure. This method ensures less dye-dependent internal variability compared to the direct labelling methods where dye incorporation is done during cDNA synthesis in one single and sequence dependent reaction. Artefacts may be a problem due to steric effects of the large fluorochromes. First, aminoally-linked cDNA is generated from mRNA with the aid from a dT oligo primer, aminoallyl-dUTP, remaining dNTPs and a reverse transcriptase, before the dyes are coupled to the aminoallyl-linked cDNA in the second step. During this work, we have used the FairPlay™ labelling kit from Stratagene, where a longer centrifugation time at the cDNA purification step was the only alteration to the instructions form the manufacturer. All reagents and columns in this kit were otherwise used without exceptions. For additional information, please refer to the instruction manual from the supplier or (Schena et al., 1995).

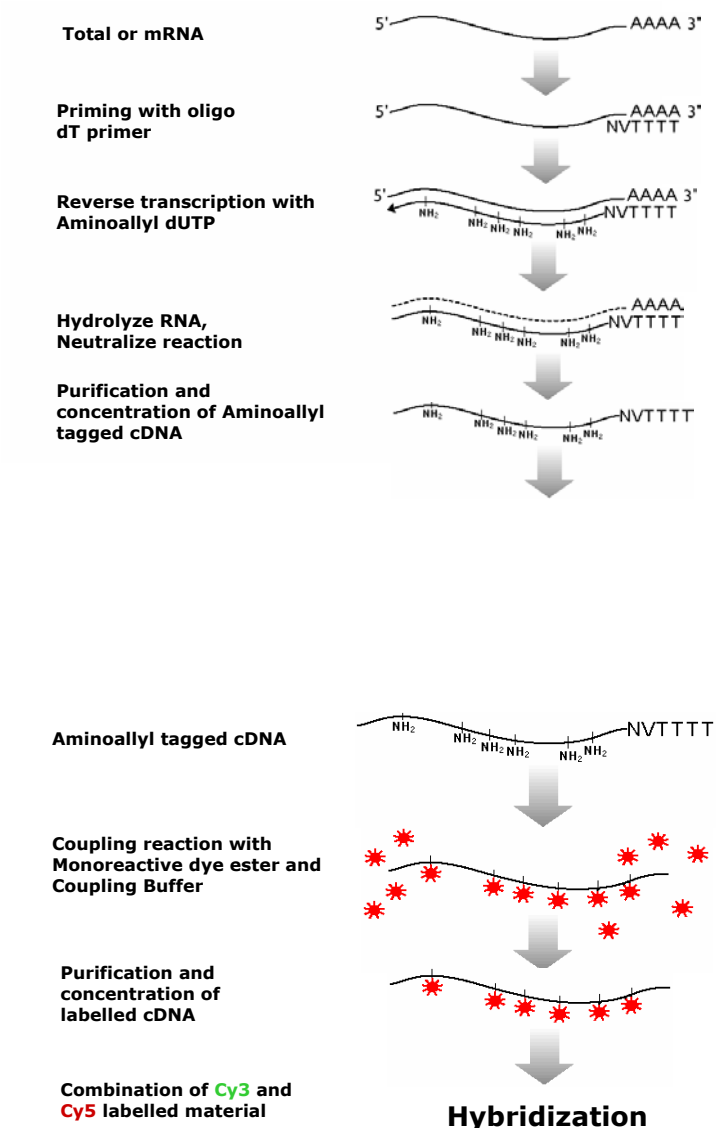


Figure 2.5: Illustration of the indirect aminolabelling technique. Figure courtesy of Ms. Signe Indahl, the Microarray Core Facility, DNR.

Additional reagents:

- Human Cot-1 DNA (placental DNA enriched for repetitive Alu/Kpn sequences) used to suppress repetitive Alu/Kpn sequence cross-hybridisation.
- PolydA used to bind the polydT tails of the generated sample/reference cDNA and probes (Invitrogen), to suppress polydT-hybridisation to A-rich sequences.

- Hybridisation buffer #3 from Ambion, containing 20% formamide, SSC and t-RNA.

Additional solutions, buffers and reagents were made according to the Fairplay protocol.

2.7.2 cDNA hybridisations with the GeneTac hybridisation station

After spin column purification, the labelled reference and sample were mixed before addition of Cot-1. Following SpeedVac-assisted up-concentration of the labelled probes, (vacuum centrifugation), they were mixed with polydA and hybridisation buffer to a final volume of 112 µl. 3 minutes at 100 °C and a 15 minute incubation at 42 °C enables Cot-1 to bind to repetitive sequences and thereby prevent cross hybridisation. During this time, the slides and equipment had been washed and prepared for loading. Samples were loaded and hybridized in a stepdown procedure from 65 to 50 °C over night. *The cDNA microarrays have cDNA fragments from the IMAGE consortium deposited onto the surface of a GAPS (gamma amino propyl silane) coated glass slide, 15000 spots in total. All slides were printed at the Radium Hospital Micro Array Core Facility (uio.mikromatrise.no). All hybridisations took place at the Core Facility.*

2.7.3 cDNA Wash

Stepdown from 50 to 42 °C, using three solutions of increasing stringency.

[Wash1 = 2 x SSC, 0,1% SDS, Wash2 = 0,1 x SSC, 0,05% SDS, Wash 3 = 0,1x SSC]

The slides were finally washed by vigorous shaking in 0,05 x SSC (2 x 2 min) before they were spun dry for 3 minutes at 1000 rpm in a swingout centrifuge (Eppendorf 5810 R).

2.7.4 Generation of image and raw data

In this work, a high resolution-scanner from Agilent enabled early visualisation and creation of a result file from each array slide. After scans at 635 (Cy5) and 532 (Cy3) nm, the GenePix software version 4.0 (www.axon.com) was applied in this work to give result files and TIFF pictures that were uploaded to and further treated in the BASE software platform.

A "GenePix Array List" file, (GAL-file), further links spot pictures to their respective reporter (i.e. gene) ID through the gridding procedure. When combined, the merged TIFF image together with the GAL-file permits positioning of the spots before ratio analysis. In GenePix, the user can "flag", i.e. manually tag and hence either remove poor quality spots, or keep those that have not been detected when TIFF and GAL were combined. A spot affected by dust and fingerprints et cetera may thus be removed. A GenePix Result file (gpr file) will further be generated. The gpr file and TIFF images constitute the raw microarray data set. In this work, the raw data was uploaded to the BASE-environment for filtering and normalisation.

2.7.5 Analysis of microarray data in the BASE-environment

BASE (at http://alba.uio.no/base) is an open network created by Saal et al., the University of Lund, Sweden (Saal et al., 2002). The software enables uploading of scanned pictures and their respective result files, to create data sets and gather experiments. BASE has several algorithms, enabling filtering and various methods for computer data normalisation, clustering and plotting of data.

Microarray data acquired in this work was processed in BASE following a series of steps. First, both sample and reference TIFF images plus the adherent result data file were combined in a single hybridisation experiment. In general for each spot, the sample signal divided by reference signal creates a signal ratio representing the direct amount of mRNA found in the sample. For representation purposes, this ratio is often log2-transformed. The raw data can be used in several experiments, and the experiments may be divided into so-called bioassays. During this correlation process, the background intensities are subtracted from the spot intensities: mean intensity in the spot foreground, and median intensity in the background for lowered bias. (Median background gives contamination and dust particles less significance.) Further on, when discussing BASE, the term "channel" refers to either reference or sample data.

Filtering raw data

Signal to noise

So-called doughnut (a spot lacking signal in its centre), may be negatively affected by the signal to noise (SNR)-filtration step, where spots with a signal-to-noise ratio of less than a fixed value (SNR mean or median in both channels lower than a fixed value), are removed. This step may be avoided if the array has a number of doughnuts, because the software may incorrectly remove such spots.

Ratio controls

Commercially available ratio controls consisting of DNA from various organisms give the user the possibility to control Cy-dye incorporation, non-specific hybridisation and the signal dynamics of each experiment. Scaled amounts may be printed on the microarray slides so that fixed ratios of e.g. 10, 3 and 1 may be obtained for each channel. In addition, negative controls (e.g. water or probe print buffer only) may be added to the microarrays to adress any non-specific hybridisation. These ratio- and negative controls are printed on the array, and RNA that will hybridise to these spots may be included at the cDNA target preparation stage. Naturally, the amount of added RNA at the cDNA target preparation step depends on pipetting accuracy, thus perhaps giving a false impression upon examination in the scanner software. Pictures may indicate that signal ratios are evenly spread, even though this may be caused by too much added control RNA to the sample RNA at the cDNA preparation stage. Taken together, these ratios may be used to assay dye incorporation, unspecific binding and the dynamics of the hybridisation (i.e. diversity of signal intensity from 0 to the maximum of 65000), but they should be used with care. In this work, the Lucidea scorecard ratio controls from Amersham Biosciences were used, and negative controls were applied for non-specific hybridisation control. They were not, however, used for normalisation purposes. In BASE, the scorecard spots, i.e. ratio-, dynamic range- and negative controls were excluded from the data set before normalisation. In addition, irregular spots with a small diameter may be removed. In this work, spots with a diameter of less than 60 μm were excluded.

Figure 2.6: Illustration of the Lucidea scorecard ratio control used in this work. Red spots are upregulated, and green signifies downregulated spots, whilst yellow spots have unaltered expression values when compared to the reference.

Removal of spots that cannot be log2-transformed

A second filter, here the "DNR-filter", may be applied to remove all spots with intensities below a set threshold. This filtration is performed because spots with intensities close to zero in one channel (and high values in channel 2) may obtain huge and inaccurate ratios after transformation, and may thus be efficiently replaced with low value to "stabilise" the ratios. The filter prepares the data for normalisation by removing intensity values that cannot be log2-transformed (e.g. intensities ≤0), i.e. spots that may not be correctly normalized. The "DNR-filter" can be adjusted, but the standard setting filters out a spot that has intensity lower than 2 times its standard deviation (SD) in either of the two channels from the data set. The standard filter keeps spots that fail in one channel, replacing the failed with the value 0.

Normalisation of raw data

After filtering, the data must be normalised, i.e. adjusted to adjust the Σintensities in channel 1 equal to the Σintensitites in channel 2. Normalisation corrects the data set for uneven amounts of RNA in each channel and uneven incorporation of the two Cy-dyes. Most normalisation methods assume that without any experimental variability, the majority of the genes in the data set have an average expression ratio of 1. In general, a normalisation factor is calculated and applied to the entire data set. One such normalisation method is the "global normalisation method", where all spots are taken equally into account. The intensities are scaled in such a way that their geometric mean (square root (intensity1 * intensity2)) is kept constant but the median of their ratios (i1/i2) is shifted to 1. More precisely described by Yang et al. (2002), one assumes that the red and green intensities are related by a constant factor; i.e. **Red= k*Green.** *The center of the distribution of log ratios is shifted towards zero:*

$\log_2$ [Red/Green] → $\log_2$ [Red/Green]-constant = $\log_2$ [Red/(k*Green)]

The constant = log_2 [k] is most often chosen as the mean or median of the intensity log ratios M for a data set. However, if the scanned arrays harbour non-linearity (banana shape) due to e.g.unequal background in the two channels, or if the signals are outside the linear range of the scanner, this calls for the 'locally weighted regression' (lowess) normalisation method. However, before a normalisation method is chosen, is is common to plot the M (log_2 ratio) values versus the A log_{10} [sqrt (Ch1xCh2)] values for each experiment. This plot depicts filtered raw signals and their symmetry around the horizontal x-axis, telling the user whether 1) the ratios are intensity-independent, and 2) the mean ratio is on average 1 (0 on a log scale) and 3) the number of up-and down-regulated genes are equal. It provides segment by segment intensity-based normalization based upon an MA plot (figure 2.7). A user defined factor *f* describes the fraction of data that should be used to smoothe the plot at each segment; the larger the value, the smoother the fit. The above formula for global normalisation is altered as follows: **log_2 [Red/Green] → log_2 [Red/Green]-constant(A) = log_2 [Red/(k*(A)*Green)]**

The constant(A) is the Lowess fit to the MA plot.

The major difference between the two methods can be said to be the indifference of lowess to a small group of differentially expressed genes that would have a large impact in the global method, where all points are valued as a whole. Depending on the width of the segment used for lowess normalisation, it is important to delimit the use of lowess on an otherwise "healthy" data set, as this method may "repair" a good spot (Yang et al., 2002). After normalisation, the data may be exported for cluster analysis, pathway mapping and statistical analyses, options not used in this work. For additional information, please refer to the results and appendix sections.

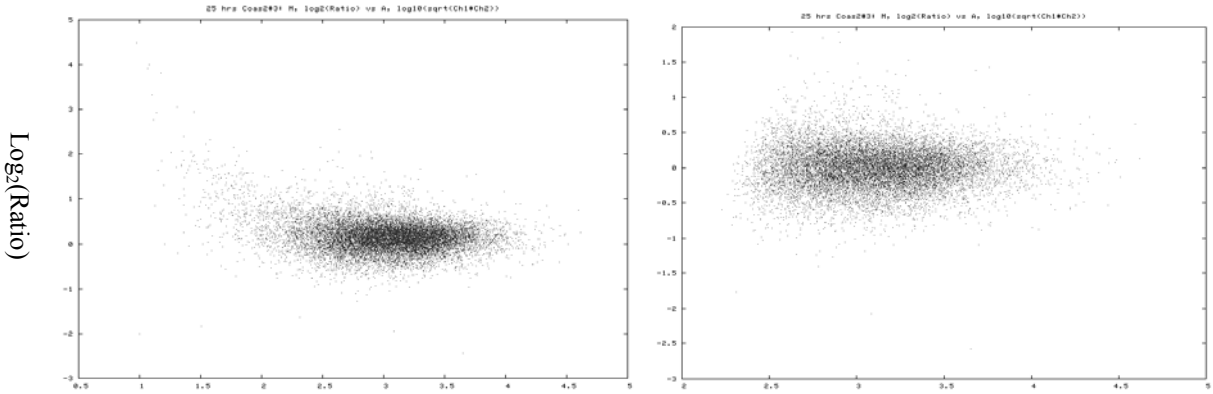


Figure 2.7: MA plot depicting the relationship between gene expression ratio and intensity on a log-log scale for the 25 hour sample taken from the time series used in this work. Data have been plotted before (left) and after normalisation (right). This positions the mean ratios around 0 (log scale).

2.8 Southern Blotting

Southern blotting is comparable to northern and western blotting techniques, but used to identify and detect specific deoxyribonucleotide sequences in genomic DNA. Through this method, information regarding whole genes and gene fragments may be gathered for DNA profiling and comparison to expression data, e.g. tumour profiling at the genomic, transcribed and translated stages (Southern, 1975).

2.8.1 Isolation of genomic DNA with the DNAZol reagent

The samples were all submitted to genomic DNA purification using the DNAZol reagent, (Invitrogen). The reagent contains mostly Guanidine hydrochloride, a strong protein denaturant, comparable to the phenol-based Trizol used for total RNA isolation. The cells are lysed in DNAZol, removing proteins and cell constituents, and total genomic DNA will then be precipitated along with total RNA using 100 % ethanol. After a brief 75 %-ethanol wash, the pellet is "fished" out of the solution with a pipette tip, and briefly air-dried before it is resuspended in 8 mM NaOH, pH 12. The alkaline NaOH serves two purposes; both increasing the rate of DNA dissociation compared to that at lower pH, as well as hydrolysing the remnants of total RNA still present in the sample. Quantification of the genomic DNA is performed after removal of total RNA in the sample (for quicker removal of RNA, RNAse A may be applied after ethanol-mediated precipitation). The pH is subsequently adjusted using

0,1 M HEPES-buffer. In this work, DNA was at first quantified as described in section 2.1.3, and secondly verified by 0,8% agarose gel electrophoresis with EtBr for visual control. (Genomic DNA often dissolves slowly, and is hence difficult to quantify properly.)

2.8.2 Enzymatic digestion of genomic DNA and electrophoresis of digested products

The genomic DNA can be digested with restriction enzymes, creating a smear on an agarose gel after overnight digestion, due to the complexity of the fragments. The DNA may further be blotted onto a positively charged nylon filter, which subsequently can be hybridised to radioactively labelled probes. In this work, Church hybridisation was used equally for northern and Southern hybridisation. A standard enzymatic digestion with λDNA was used as positive control:

Genomic DNA and STD genomic normal DNA	7 µg
10 x restriction enzyme buffer	6,5 µl
Spermidine (carrier protein) final 4 mM	2,5 µl
BSA if necessary, final 0,1 mg/ml	6,5 µl
Enzyme (20 U/µl)	2,5 µl
Sterile water to:	65 µl

- Remove 1/10 of the total volume, then add 200 ng of λ DNA. Incubate separately.
- Make a positive control digest with the enzyme, 200 ng λ DNA to 8 units of enzyme.
- Incubate over night at the enzyme's indicated working temperature.
- The next day, cast two 0,8% TAE agarose gels with 10 µg/ml final EtBr.

The digest controls will give feedback on the reaction's rate of completion. If the controls show that λ DNA has been fully digested, one can proceed by adjusting sample volume upon visual comparison to the intensity of normal DNA of known concentration. The remaining 60 µl should be loaded onto the second 0,8% gel, and samples should be run at 40 V for 3-5 hours. A molecular weight standard marker may be included at this point.

In this work, *Hind*III was used in parallel with a set of 6 samples, plus control digests as outlined above.

2.8.3 Blotting from gel to filter membrane

At the visualization step, the gel may be left under UV illumination for 1-2 minutes for further nicking of the DNA and thus optimization of transfer from gel to membrane. The gel is further denatured for 30-45 minutes in an alkaline buffer with gentle shaking (0,4 M NaOH, 1,5 M NaCl), and the blotting apparatus is mounted as illustrated in figure 2.7, using a standard Southern blot transfer buffer (0,4 M NaOH, 1,5 M NaCl). After incubation over night, the gel is removed from the blotting apparatus, and the positions of the wells are located with a soft pencil on the nylon filter. A soak in 50 mM NaH_2PO_4 pH 7.2 for 30-45 minutes will neutralise the alkaline transfer buffer. The filter may further be treated as described for a northern blot (although all Church solutions intended for DNA use are made with regular ddH_2O, and not DEPC-treated water).

2.8.4 Probe preparation

Hygromycin ORF

Specific primers were designed in silico towards a subsection of the hygromycin open reading frame. With the pcDNA5FRT plasmid as template, a standard PCR was performed (annealing at 53°C), using a Pfu polymerase in four 100 µl reactions. Of the total 400 µl, 200 µl were submitted to PCR and gel purification according to 2.3.3 and 2.2.4.

lacZ-Zeo fusion gene

6 µg of the pFRT*lac*Zeo vector was submitted to enzymatic digestion with 35 units *Eco*RV and 40 units *Sac*I for two hours at 37°C to yield a short 807 bp *lacZ*-Zeo fragment, which was visualised after electrophoresis, and later gel purified as described in section 2.2.4.

The Southern technique was applied to various DNA samples from the Flp-In system, please see the results section for further information.

2.9 Working with mammalian cells

All cell work was performed under sterile conditions in a laminar air flow (LAF) bench. Most of the work was done on adherent Flp-In™ 293 cells (the parental 293 cell line has the ATCC-number CRL 1573), although other cell lines were cultured during shorter periods of time. Cells were kept in humidified incubators (water vapour), where the CO_2 and O_2-levels were fixed at 5%. For medium recipes, please refer to the appendix section.

2.9.1 The Flp-In system

The Flp-In™ system allows integration of a gene of interest in eukaryotic cells at a specific genomic location (Craig, 1988); (Sauer, 1994). A cell line harbouring a single 34 base pair Flp Recombination Target (FRT) site enables the creation of stably transfected cells. Two plasmid vectors are co-transfected into the FRT-containing cells to generate isogenic, but non clonal cell lines expressing the inserted gene of interest from the FRT locus, as initially described by O'Gorman et al. (1991). The "Flp" recombination system has been derived from the budding yeast *Saccharomyces cerevisae*, and selection of positive clones is based upon addition of the antibiotic Hygromycin to the cell growth medium. Two 293 cell lines (with only one FRT site each) were used: The Flp-In 293 cells offer constitutive expression of the inserted gene, whereas the Flp-In T-Rex 293 cells carry a TetOn promoter which blocks transcription in the absence of tetracycline. In each parental cell line, the lacZ-Zeocin-resistance fusion gene is expressed, making them resistant to the antibiotic Zeocin. The T-Rex cell line harbours a second resistance trait, having the pTR6© plasmid which codes for both Blasticidin resistance and the essential TetRepressor protein.

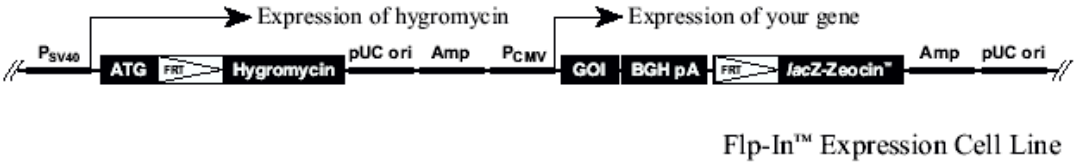


Figure 2.8: Illustration of a Flp-In expression cell line depicting the integration site after Flp-recombinase assisted insertion of the pcDNA5FRT plasmid. GOI = gene of interest.

Two plasmids were co-transfected to create each of our transfected cell lines:

-The **pOG44** vector constitutively expresses the Flp recombinase (Broach et al., 1982), under the control of a human CMV promoter. The remaining FRT or FRT/To vector, (one for each cell line), contains the gene of interest in frame with the coding sequence for the antibiotic hygromycin and the unique 34 bp FRT site, all under CMV promoter control. Upon Flp recombinase-mediated integration of the second plasmid into the genomic FRT site, hygromycin resistance and Zeocin sensitivity is conferred to the cell line as the lacZ-Zeocin reading frame is disrupted by the **pcDNA5FRT/pcDNA5FRT/To** plasmid. Upon transfection of these cells, the isogenic nature of the transfected cells, i.e. proper insertion of the gene-containing plasmid, can be verified through assays for Zeocin sensitivity and/or β-galactosidase staining. In this work, we chose to verify Zeocin sensitivity, learning more about the Flp-In system (for additional information, please see the results section). Eight kinds of cell lines were generated and used during this work. The pcDNA5FRT plasmid harbouring the gene of interest was transfected into the Flp-In 293 cell lines for constitutive expression, as the pcDNA5FRT/To with inserted gene into the Flp-In T-Rex 293 cells for tetracycline-inducable expression.

Table 2.6: Constructs created for use with the Flp-in system.

Constitutive expression	**Tet-inducable expression**	**Produced by**
pcDNA5FRT/COAS2	pcDNA5FRT/To/COAS2	Mr Magne Skårn
pcDNA5FRT/COAS2MycHis	pcDNA5FRT/To/COAS2MycHis	Mr Erik B.Paulsen
pcDNA5FRT/egfp		Mr Jørn Henriksen
	pcDNA5FRT/To/egfp	Mr Magne Skårn
	pcDNA5FRT/egfp/COAS2	Mr Magne Skårn
	pcDNA5FRT/COAS2/egfp	Mr Magne Skårn

The egfp coding sequence was subcloned from the pEGFP-N1 vector (Clontech). COAS2-vectors created in this work all harbour the COAS2 492 nucleotide ORF, courtesy of Dr. Leonardo A. Meza-Zepeda and Professor Dr. philos Ola Myklebost.

Creation of *COAS2*MYC-His plasmids

The C-terminal 9E10Myc-His_6 tag was subcloned from the pcDNAMYCHis B(+) plasmid (Invitrogen), using two different primer pairs. First, *COAS2* was amplified from the pcDNA5FRT/To/*COAS2* plasmid using the COAS2XbaIUpper and COAS2SfUILower primers in a standard 20-cycle PCR, followed by gel purification. Both the pcDNA3.1MYCHis B (+) plasmid and the purified *COAS2* was digested in a single reaction with *XbaI and SfUI.* The fragments were ligated, bacteria were transformed with the recombinant plasmids and PCR screening applied to identiy positive clones. After plasmid purification and verification by enzymatic digestion and MEGA BACE sequencing, the second PCR with HisTagCOAS2StopXhoI and COAS2BamHI primers produced a COAS2-MYC-His duplex DNA which was gel purified and treated exactly as the above PCR product for creation of the pcDNA5FRT/COAS2MYCHis and pcDNA5FRT/To/COAS2MYC-His plasmids. A flow chart for the entire procedure is shown in the appendix section. Proper induction of four of the *COAS2* cell lines was confirmed in a northern blot with a *COAS2* cDNA restriction digest as probe. The Myc-His tag was not selectively confirmed; please see the result section for data and thoughts on the subject. Both the *egfp*-transfected and the *egfp/COAS2* fusion construct-transfected cells were examined for egfp expression at 488 nm with a Zeiss inverted fluorescence microscope. Please refer to the results and appendix sections for illustrations of the cloning strategies and characterisation of the Flp-In transformants.

2.9.2 Growth and harvest of mammalian cells

Cells frozen in liquid nitrogen were quickly thawed and transferred to a 75 cm^2 cell flask. Gentle resuspension in 12 ml antibiotic-free medium was followed by incubation at 37°C over night. The following day, cell line-specific (i.e. with all necessary additives) medium was added after a brief wash with 37°C PBS. When the cells were approximately 80% confluent, they were washed 1 x with PBS and harvested using trypsin/EDTA. The trypsin is more efficient at 37°C, so some cell types were incubated for short periods, typically 2-3 minutes in at this temperature. *Cells attach to the plastic flask through membrane glycopeptides. Trypsin breaks a polypeptide chain at the C-terminal end of the amino acids arginine and lysine, and thus cleaves glycopeptidic bonds, detaching the cells from the flask. Serum protease inhibitors plus Ca^{2+}/Mg^{2+} inhibit trypsin, explaining partly why a PBS-wash is necessary when trypsin*

is to be used. The chaotrophic EDTA is used to strengthen the effect of trypsin as it binds free calcium ions in solution (Kielberg, 1993).

If downstream purposes such as cell counting or freezing requested so, the cells were transferred to a 10 ml polystyrene tube, sealed in the sterile bench and centrifuged for 5-6 minutes at 200 x g = 1000 rpm in a centrifuge at room temperature (Heraeus Megafuge 1.0 or Labofuge 400R). The excess liquid was removed, and the pellet was either washed once in PBS, or resuspended in fresh medium before an eventual count of the cells was performed. All cells were subcultured differently, according to cell type and growth rate.

2.9.3 Counting cells

The cells were all counted using a hemacytometer, or cell counting chamber (Bürker, 0,1 mm x 0,0025 mm^2). This unit contains two rectangular cavities, made up by nine major squares each one again divided into sixteen squares. To count, the cells are trypsinated as in 2.9.2, transferred to a polypropylene tube and centrifuged at 1000 rpm for 5 minutes. The pellet is further resuspended in a known amount of fresh medium. A small aliquot is extracted with a pipette, and 10 µl are applied to the cavity between a 10 mm^2 cover glass and the hemacytometer. The amount of cells per ml is thereafter determined in a light microscope by counting several squares of known volume and extracting the mean value, which is multiplied by a chamber-dependent size factor, in this case 10000. The result is given as # cells per ml. For an illustration, please see (Kielberg, 1993).

2.9.4 Freezing cells

Cells were harvested as described in section 2.9.2, and the pellet resuspended directly into ice cold DMEM with 20% FCS, 10% dimethylsulphoxide (DMSO), 2 mM Glutamax™ and aliquoted into 1,5 ml cryovials (Sarstedt). The tubes were transferred to a plastic container with isopropanol and left at -70°C for 24 hours before they were frozen and stored in liquid nitrogen. The container mentioned above permits a one-degree per hour steady decrease from 0 to –70 degrees, which minimizes the loss of cells.

2.9.5 Screening for mycoplasma contamination in cultured cells

The contamination of cells in culture by mycoplasma and other members of the class *Mollicutes* is a common occurrence. The metabolism of the attacked cells changes towards the non-predictable, making experiments unreliable and non-reproducable. A simple kit from ATCC has been used for the detection of mycoplasma in cell cultures during this work. It is based on nested PCR, running 2 PCRs sequentially and finally visualising eventual bands with 2% agarose electrophoresis. Please refer to the supplier's homepage at http://www.atcc.org for a detailed protocol (Cat.nr 90-1001K, version 2.0).

2.9.6 Transient and stable transfection of mammalian cells

There are several ways to *transfect* a eukaryotic cell, of which the author has used two techniques. The transfer of biomolecules across the cell membrane demands transient permeabilization or assisted transport across the hydrophobic region. Electroporation exemplifies the first method, where the cell membrane is perforated after an electric pulse has been submitted to the sample for some milliseconds, permitting exogenous DNA to enter. In this study, electroporation was used for transient transfection only, i.e. introduction of the recombinant DNA into the recipient cell line without its integration in the cell line's genome.

Standard protocol for electroporation of mammalian cells

This protocol can be found as protocol # 304 at the BTX Division of Genetronics homepage. The BTX Electro Cell Manipulator™ electroporation device with 4 mm gap cuvettes was used. A cell suspension of 2,5-4,0 x 10^6 cells/ml growth medium (total volume: 400 µl), was transferred to the cold P/N 640 disposable cuvettes before addition of the plasmid DNA to be transfected. After the 26-45 ms pulse, the cuvettes were kept on ice until transfer of the cells to 75 cm^2 cell flasks supplied with fresh growth medium (Baum et al., 1994); (Sambrook, 2001). Transfection efficiency is influenced by a number of factors; as the length of the pulse, the capacitance value, DNA concentration as well as the strength of the applied electric field and ionic composition of the medium. The apparatus has only one adjustable variable, namely the voltage. We therefore kept the amount of DNA and medium composition fixed, and rather adjusted the voltage and thereby the pulse length to optimise the transfections. Samples were kept on ice at all times, even though some favour room temperature after the electroporation.

Ideally, protocol # 151 from BTX with an apparatus such as the ECM 600/630 or equivalent should be used for the transfection of human 293 cells. A more gentle way to transfect eukaryotic cells is by *liposome*- or non-liposome based transfection. A liposome will surround the hydrophilic sample to be transfected, and helps it transfer across the cell membrane and into the cytoplasm. Non-liposomal reagents will in similar manners assist the translocation across the phospholipid bilayer. This class of transfection reagents is less harmful to the cells, and their potency maximizes the odds for further transport into the nucleus, a matter of importance when stable transfection assays are performed.

Transfection of mammalian cells with FuGene6

The FuGene6 high efficiency transfection reagent is a non-liposomal, proprietary blend of lipids with low or no sign of cytotoxic behaviour. It was used for all transfections in this work. After optimisation of the conditions (see table 2.7), we performed all transfection experiments according to the standard protocol using serum-free DMEM cell growth medium mixed with (1-2,5) µg of DNA and (3-12) µl of FuGene6 in a sterile eppendorf tube at room temperature. After 15-45 minutes of incubation, the mix was added dropwise to the adherent cells, which had been seeded out 24 hours earlier in a 35 mm dish using 1-3 x 10^5 cells in total. The amount of FuGene was always kept higher than the amount of DNA, and egfp-expressing plasmids were used as controls in all types of transfections. The cells were grown in antibiotic-free medium until selection, a timespace of 24-48 hours. For a more detailed and updated protocol, please see http://www.roche-applied-science.com. As indicated further on, optimisation was performed using transient transfection and flow cytometry measurements, indicating an ideal ratio of 6:1 for µl FuGene6 vs µg plasmid DNA, although 3:0,5 also has been found to be well suited. This was transient, however, and transfections very much depend on the purity of the molecules to be transfected. Co-transfection increases the complexity, and for stable transfections in the Flp-In system, we always co-transfected the pOG44/pcDNA5-plasmids in a *minimum ratio of 9:1*. After some experiments, we found that the amount of DNA is to be kept low to avoid non-specific stable integration using the Flp-In system. Please see the results and discussion sections for further information regarding the Flp-In system performance.

Induction of 293 Flp-In transfectants with tetracycline

Tetracycline was added to the growth medium to a final 1 μg/ml for induction. Tetracycline will bind to the Tet-Repressor protein, which causes a change of protein conformation and thus liberation of the CMV promoter (please see the next section for additional information). When assaying for *COAS2* expression with northern blotting, we seeded out 2-3 x 10^6 cells in 75 cm^2 flasks the day before induction. For microarray experiments, we used 175 cm^2 flasks and seeded out 3,75x10^6 cells in these the day before induction, yielding 110-360 μg of RNA, respectively.

Testing 293 transformants for Zeocin and Blasticidin sensitivity

To assay the function of both Zeocin and Blasticidin, we performed experiments throughout the work period. Zeocin was given at 5-1000 μg/ml for clone verification experiments, and 10 μg/ml Blasticidin was fed parental cells to assay for homogenicity in the cell population. Zeocin has also been given to LOX cells (kindly provided by Mrs. H. Høifødt of the Dept. of Tumor Biology), to verify drug function efficacy.

2.9.7 Creation of *COAS2*-xenografts

The FRT/To/*COAS2 clone#3,* FRT/*COAS2 clone#1*, and the parental Flp-In 293 cell line were all grown without antibiotics for 2 passages, and harvested and counted according to sections 2.9.2-3. Cell densities: FRT *clone#1= 3,2 x 10^6/ml; FRT/TO clone#3=6 x 10^6/ml; Flp-In 293=8,5 x 10^6/ml.* 2 x 0,2 ml was injected subcutaneously into immunodeficient nude mice by staff at the Animal Research Section at the Norwegian Radium Hospital. When tumours had reached roughly 6 mm in diameter, they were split and pieces of Ø = 2mm transferred to new mice. After 2 passages, the inducible FRT/To/*COAS2 clone#3* was submitted to an induction experiment initially described by (Kistner et al., 1996), where the tetracycline analogue doxycycline (which has a $t_{1/2}$ two times the one of tetracycline), was fed at 200 μg/ml to the mice in drinking water with 2% sucrose to mask the bitter taste. Mr. Christoph Müller performed the experiments. Tumors from early passages were stored for total RNA purification, immunohistochemistry and other downstream applications.

2.10 The SRB protein assay

The SulphoRhodamine B dye binds proteins and is used for in vitro chemosensitivity testing of cell lines. The assay is rapid, inexpensive and sensitive, permitting at its most primitive the measurement of doubling time for cell lines. Under mildly acidic conditions, SRB binds to basic amio acid residues, providing a sensitive index of cellular protein content that is linear over a cell density range of at least 2 orders of magnitude. The assay was performed essentially according to (Skehan et al., 1990). Dilutions of cells are seeded out in four 96-well plates and grown over night, ranging from 1-5000 cells per well (6 parallels for each dilution). At t_0, cytotoxins may be added. For measurement of doubling times, the cells are merely fixed, washed and air-dried after 24, 48, 72 and 96 hours, respectively. 50% (w/v) trichloric acid (TCA) was applied to fix the adherent cells in the SRB assay, but several agents are currently in use[5]. When MeOH is used, permeabilisation of the cell membrane may cause a leakage of internal soluble proteins towards the medium. After fixation, the cells are rinsed in lukewarm tap water, shaken and left to dry at room temperature. When all plates have been collected, they are submitted to SRB staining[6], washed in 1% CH_3CO_2H (w/v), shaken and air-dried, and finally resuspended in 10 mM Tris pH 7.5. A plate-reading spectrophotometer reads the intensities at 540 or 490 nm (absorbance maximum and half of the latter, respectively). There are numerous downstream applications of this assay. In this work, it was used to create growth curves plus estimating cell culture division time and also for preliminary chemosensitivity trials. The above experiments were made possible with the help of Mr. Paul Noordhuis. Mr. Christoph Müller performed cytostatica-sensitivity assays using the *COAS2*-transfected cells.

Calculating doubling time

The doubling time was calculated using the following formula, where A represents the absorbance at 540 nm:

$$\text{Doubling time (hours)} = \text{Experiment total time (hours)} \times \frac{\log_{10}(2)}{\text{growth ratio}}$$

Where the growth ratio = $\log_{10}\frac{A(t)}{A(t_0)}$, where t = time of harvest and t_0 = 0 hrs

[5] Formaldehydes cross-bind proteins in the sample, while alcohols (e.g. methanol, MeOH) fix cell membranes by dissolving lipids and denaturate proteins, forcing the latter to aggregate.
[6] 0,4% SRB (w/v) in 1% CH_3CO_2H (w/v)

2.11 Cytostatica sensitivity assays

Cells with constitutive expression of COAS2 and the parental 293 Flp-In cell line were plated as indicated earlier in the text. Methotrexate (MTT), Cisplatin (CDDP) and Doxorubicin (dox) were added to the growth medium at day 1, and cells were further treated and counted as indicated above. The drugs were administred in the following doses: 0,01-100 mM (MTT), and 0,01-1000 nM (CDDP and dox).

2.12 The Courtenay-Mills' soft agar assay

Cells' ability to grow without anchorage, or their *clonogeneicity*, is tested using this simple growth assay. This ability is correlated with tumourigenicity (Courtenay and Mills, 1978).

Standard protocol

Two ml blood from one nude rat is mixed with 5 ml phosphate buffered saline (PBS) without neither Mg^{2+} nor Ca^{2+}. Centrifugation at 2800 rpm for 20 minutes is followed by removal of the supernatant and intersection, the latter consisting of unwanted leukocytes. After addition of 5 ml PBS, and mixing by inversion, the step is repeated once. PBS is added to the initial blood volume, and remaining immunogenic factors in the nude rat blood are inactivated at 42°C for 60 minutes before use. (This solution will provide factors that promote growth of the cells in soft agar.) The prepared growth medium must contain 15% inactivated Fetal Calf Serum (FCS), 2mM L-Glutamine or equivalent additives, and 2mM penicillin-streptomycine. The Fham-12 medium may in some cases be replaced with the standard growth medium if the serum amount is adjusted to 15%. Each cell suspension is created in triplicate, and a typical range for a chosen cell line would go from a final 5000 cells/ml to 50 cells per millilitre. The cells are seeded out in Fham-12 cell growth medium, mixed with erythrocytes (diluted 1:8), and sterile 5% Beetor Soft Agar solution (to final 0,5 %). Gentle mixing must be performed before the agar solidifies. Incubation takes place in a 5 % CO_2/O_2-controlled incubator for one or two weeks. (The nitrogen ensures 5 % O_2 in the incubator at all times.) After one week, some tubes are counted, and the remaining lot of cast-in cells receive fresh cell medium if prolonged incubation is found necessary. After a fortnight, the jellylike pellet is squashed between two gridded plastic surfaces; the grid eases the counting of the tumour-like cell colonies.

2.13 Flow cytometry and fluorescence

Fluorescence can be defined as the molecular absorption of light energy (photon) at one wavelength and its re-emission at another wavelength. Light-absorbing molecules are known as chromophores, whilst molecules that both absorb and emit light are known as fluorochromes or fluorophores. The process of fluorescence may be divided into three phases, i.e. excitation, the transient excited state and emission. The phenomenon can be described as follows: The incident monochromatic light excites the sample. If the incoming energy is sufficient, some electrons are lifted to a level of higher energy in the sample. The surplus energy is passed as transmitted light When the electrons return to their initial energy state, the inititally absorbed excitation energy is being released as emission light; photons of a higher wavelength. A fluorophore's extinction coefficient ε characterises its ability to absorb light energy, and each fluorophore has excitation and emission maxima related to this variable. The (energy) difference between a fluorochrome's maximal excitation and emission wavelengths is known as the Stokes shift.

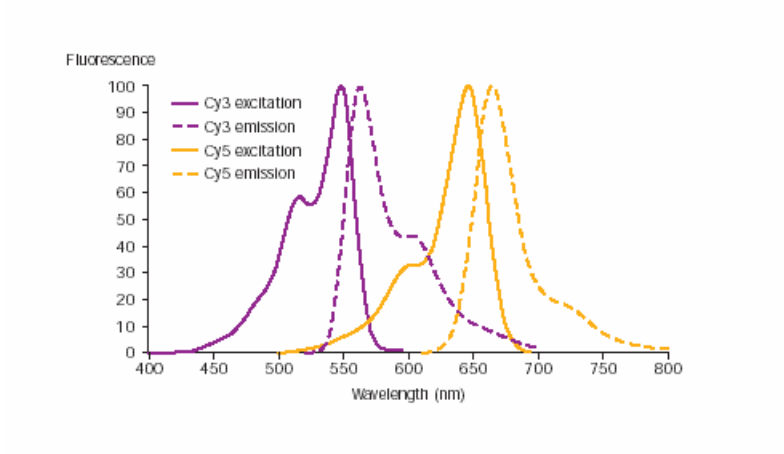


Figure 2.9: The excitation and emission spectra for the Cy3 and Cy5 dyes used in this work. The Stokes shift is the difference between the continuous- and dotted line peaks, respectively.

The fluorescence and scattered light from microscopic particles can be quantified and analyzed using the flow cytometry method. Both intracellular components and fluorochrome-tagged endogenous constituents can be visualized and quantified using this technology. Fluorescently labelled cells can be sorted and counted with a high degree of sensitivity: a few hundred molecules per cell can be detected. Endogenous pyridines and flavins are detected in addition to the extrinsic fluorescence that includes all other sorts of fluorochrome-labelled particles. In a flow cytometer, the analysed particles (most often cells) are suspended in a mobile phase and lead through a funnel-shaped tract, creating a flow that passes the internal

system laser(s). Excited by the intense light, each cell creates a brief pulse of emitted light (of lower energy thus higher wavelength) which passes beam splitters, mirrors and filters to face a photomultiplier tube; (PMT) one tube for each laser. In addition, the cells will spread the laser light, and scattered incident light is collected and measured at two angles, namely forward (orthogonally, or "forward scatter"), and at low angle ("side scatter"). The scattered light provides information on cell size, granularity and structure. All fluorescent signals are digitized in the PM tubes, and thus quantified and visualized by the aid of computer software. Sorting cells of homogenous nature in a blend of various cells can be done, but sorting of cells was not performed and will not be discussed any further.

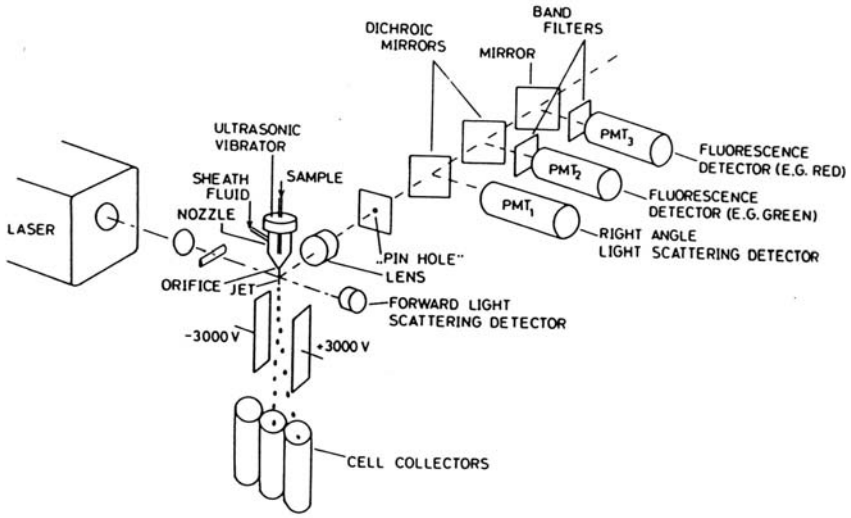


Figure 2.10: A principle outline of a laser-based flow cytometer. The flow cytometer has sorting facilities and opportunity to detect two fluorescent components (PMT_2 and PMT_3), large angel light scattering (PMT_1) and light with low angel scattering (forward light scattering detector). Figure from (Steen, 1991).

2.13.1 Optimizing transient transfection with flow cytometry

The use of green fluorescent protein, or gfp, has greatly improved molecular biotechnology (Tsien, 1998). Gfp emits in the green range of visible light when excited with blue light, and was first characterised and isolated from the jellyfish *Aequorea victoria* (Tsien, 1998). The enhanced version named egfp (Clontech) has been used in this work. For preliminary tests of transfection efficiency, we transiently transfected the EGFP-N1 plasmid from Clontech into the Flp-In 293 and Flp-In 293 T-rex cell lines with electroporation and FuGene6, respectively. Flow cytometry was used to analyze the percentage of successfully transfected living cells, with the DNA-binding dye propidium iodide (PI, see also next section) as a control marker.

(PI enters dead cells only, permitting us to remove these by gating for the viable, non-emitting cells.) Both dyes were excited with one argon laser at 488 nm, but various filters in the FACStar are applied for each dye. Egfp emits at 509 nm (green) and PI at 617 nm (red), respectively.

Sample preparation

The electroporation and FuGene6 transfection were performed as described in section 2.9.6. Electroporation of the T-Rex cell line was done with two different cell concentrations in parallel series, 3,2 and 6,4 x 10^5 cells/ml.

Transient transfection of the Flp-In 293 cell line with the FuGene6 reagent:

Table 2.7: Optimisation of transfection using the FuGene6 reagent. Initial setup of reagents.

Tube no.	1	2	3	4	5
DMEM w/o FCS (µl)	97	97	97	94	88
EGFP-N1 DNA (µg)	0,5	1	2	1	2
FUGENE6 (µL)	3	3	3	6	12

Quantification of Flp-In induction and egfp-intensity

With the stably egfp-transfected cells, we aimed to measure both percentage of transfected cells and to create an induction curve for the tetracycline-inducable clone. Cells were induced with tetracycline, and harvested at various time points according to section 2.9.2-3. Between two centrifuge runs, the cells were washed once and diluted in PBS to a final concentration of 0,5-1,0 million cells/ml. Some samples were fixed in ice-cold paraformaldehyde (final 3%) for 60 minutes on ice, spun for 5 minutes at 1000 rpm and resuspended and stored as above. Shortly before the run, the suspensions were filtered through a 50 µm mesh to obtain a maximum number of single cells for injection. Subsequently, PI was added to a final 0,5 µg/µl for detection of dead cells. With the kind aid from fellow student Lise Ramberg and Mrs. Kirsti S. Landsverk at the Flow Cytometer Core Facility at the Norwegian Radium Hospital, the analysis was performed on the Becton Dickinson

FACStar^{PLUS} (FACS = Fluorescence Activated Cell Sorter) flow cytometer and the CELLQuest™ v3.3 software.

2.13.2 Cell cycle analysis

As described in section 1.2, eukaryotic cells follow the cell cycle where the DNA content changes between phases. Estimating the cells' position in the cycle can therefore be made using a flow cytometer and a DNA-binding dye. Propidium iodide (PI, excitation maximum at 536 nm) was the chosen fluorochrome. Acting as an intercalating agent, the dye binds stoichiometrically to every 4-5 base pairs. PI also binds to RNA, necessitating treatment with ribonucleases to distinguish between DNA and RNA staining. (Ribonuclease A (RNase A) specifically attacks single-stranded RNA 3' to pyrimidine residues and cleaves the phosphate linkage to the adjacent nucleotide.) The enzyme works in the absence of cofactors and divalent cations.) To deselect or gate dead cells away from the analyzed population, PI can be used in companionship with other dyes, as it cannot enter viable cells. However, for cell cycle analysis, all dead cells are included, making the assay less sensitive.

Protocol

Two populations were seeded out at day 1: Cells with transfected COAS2 under promoter control, and parental reference cells. After induction of the cells with tetracycline, the cells were harvested and counted according to section 2.9.2-3. After one wash with PBS, the pellet was very gently vortexed and the cells were ensuite fixed in 1-2 ml of ice cold MeOH. After 1 minute of incubation, a second 5 minute spin at 1000 rpm followed before they were resuspended to a final concentration of 7,5 x 10^5 cells/ml with a PBS-solution containing 10 μg/ml propidium iodide and 100 μg/ml RNase A, the latter prepared according to (Sambrook, 2001). After 1 hour of incubation at 37°C, the sample was analyzed after filtration through a 50 μm mesh at RT using the same hardware and software as in earlier transfection experiments. The FL3 channel was used for DNA measurement as well as doublet discriminating module (DDM), with triplicates of each sample type to assess systematic bias. Samples were run at medium speed, counting a total of 5000 gated cells in each try.

A second experiment used serum-starved FRT/To/*COAS2 clone#3* cells, (fed 2% FCS only for 72 hours previous to tetracycline addition), compared to the reference cells grown in 8% FCS. Cells were induced and treated exactly as the cells in the former experiment.

2.14 Fluorescence confocal microscopy

Traditional light microscopy and fluorescence microscopy cannot produce optical sections through a 3-dimensional (3-D) specimen - e.g., an entire cell or a piece of tissue - that, to a good approximation, contain information from only one focal plane (Lichtman, 1994). In a fluorescence confocal microscope, however, this is possible, as illustrated schematically in Fig. 2.11, by moving the focal plane of the instrument step by step through the depth of the specimen, a series of optical sections can be recorded (c. f. Lichtman, 1994), seeing where the fluorochrome-tagged objects of interest are located in the given sample. Scanning the object in the xy- and z-direction (along the optical axis), allows viewing from all sides.

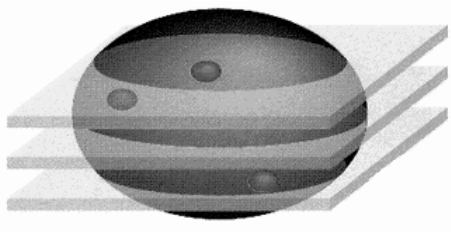


Figure 2.11: Illustration of focal XY-planes scattered at three points on the Z-axis.

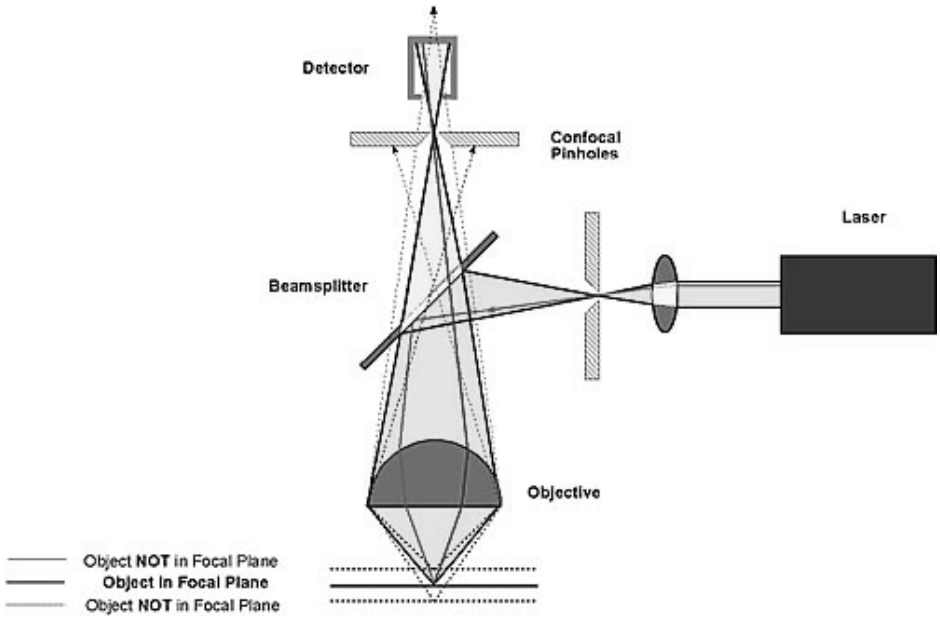


Figure 2.12: Schematic illustration of the confocal principle, courtesy of Leica Microsystems.

To image the specimen point by point, a collimated, polarized laser beam is deflected stepwise in the x- and y-direction by a scanning unit (not shown) before it is reflected by a dichroic mirror (beam splitter) so as to pass through the objective lens of the microscope, and focused onto the specimen. The emitted, longer-wavelength fluorescent light (principles of excitation and emission are described in section 2.13), collected by the objective lens passes through the dichroic mirror (transparent for the longer wavelength), and is focused into a small pinhole (i.e., the confocal aperture) to eliminate all the out-of-focus light, i.e., all light coming from regions of the specimen above or below the plane of focus. Therefore, the confocal microscope does not only provide excellent resolution within the plane of section (0.25 mm in x- and y-direction), but also yields similarly good resolution between section planes (0.3 mm in z-direction). The in-focus information of each specimen point is recorded by a light-sensitive detector (i.e.a photo-multiplier tube) positioned behind the confocal aperture, and the analog output signal is digitized and fed into a computer. At the same time, the analog photo-multiplier signal can be used to generate an image on a video monitor. To obtain a full image, the image point is moved across the specimen by mirror scanners. As for a standard fluorescence microscope, the emitted light is filtered through mirrors plus beamsplitters and collected in photomultiplier tubes (PMTs). But the difference lies mainly within the ability of a confocal microscope to suppress, i.e. lock out all light from structural sources not in focus upon image formation, as well as stray light and defocusing blurs. In this way, the image of one focal plane through a specimen is created, and by adjustment of the numerical aperture plus sensible variations of the focal plane depth, a stack of serial optic sections may be created. Such stacks may be summoned to create a composite projection image, a procedure that was used in this work with 8 serial pictures per composite image. When several fluorochromes and thus several instrument lasers are in use, their respective emission/excitation spectra may overlap, a phenomenon called “bleed through” which may be corrected upon adjustment of initial laser light strength. “Bleed through” may cause poor overall resolution, but was avoided in this work, as we only stained with one Rhodamine/TRITC-tagged secondary antibody.

Localisation experiments

First, we wanted to stain our *COAS2MYC-HIS$_6$*-transfected cells with primary Myc or His-antibodies plus Rhodamine-tagged secondary antibodies. This would enable us to localise COAS2 in the 293 cells upon over-expression.

During her work as a MSc-student, Ms. Birgitte Lygren indicated a possible interaction between the proteins filamin type A/C and COAS2. So in a second confocal microscopy experiment, we aimed to compare filamin patterns in cells with and without overexpression of COAS2. This was done with primary chicken α-goat anti-filamin (SIGMA) and donkey anti-goat Rhodamin-tagged secondary antibody (#522, a kind gift from Dr. H. Stenmark). Primary diluted 1:40, secondary 1:100. We also searched for alterations in the cells' actin pattern, using the actin-binding reagent phalloidin in a 1:750 dilution. This fluorochrome has no need for secondary antibody.

2.14.1 Standard protocol for immunofluorescence microscopy of attached cells

Used for both confocal and regular fluorescence microscopes. Modified from a standard protocol described by Dr. H. Stenmark.

Cells are seeded out on 10 mm^2 cover glasses 24 hours prior to fixation. Cells in log phase growth are washed three times with ice cold PBS before they are covered by freshly made 3% paraformaldehyde[7] on ice. After 15 minutes, the excess aldehyde groups are quenched using 50mM NH_4Cl for 15 minutes on ice. Incubation on ice for 5 minutes in 0,1-1% Triton X-100 in PBS permeabilizes the membrane, permitting the antibodies to enter the cells. 20 µl of primary antibody solution (5-10 µg/ml) are put onto a small piece of PARAFILM™ in a humidified chamber. The cover glass is then put with cells facing down towards the liquid using a pointed forceps. Incubation lasts for 15-40 minutes, and is followed by 3x5 minute washes in PBS containing 0,05 % saponin (a detergent originating from the bark of the Quillaja tree). The fluorochrome-conjugated secondary antibody incubation and 2x wash are

[7] *Fixation with paraformaldehyde preserves cell structure better than organic solvents such as MeOH (Sambrook, J., Russell D.W. (2001). Molecular Cloning, A laboratory manual, Cold Spring Harbor Laboratory Press.. Proteins are cross-linked in the simple reaction, but it does not allow access of the antibody to the interior of the specimen. This is achieved by non-ionic detergent permeabilisation on ice.*

performed as listed above, before the glasses are washed 5 minutes in PBS, soaked in water and mounted on an object glass with 3 µl of Mowiol (cells facing down.) The sample is now ready for inspection after 30 minutes at 37°C to fix the Mowiol-preparation. Cells were visualized using a Leica NT confocal microscope at the Department of Biochemistry, the Norwegian Radium Hospital.

2.15 Protein expression in bacteria and insect cells

The transition from nucleic acids to peptides and proteins implies a major increase of complexity. From a linear amino acid sequence, a three-dimensional peptide or protein will form. To investigate whether *COAS2* really gave a protein product, we launched into the vast field of protein expression and purification.

2.15.1 Prokaryotic COAS2-expression in the Ek/LIC system

The bacterial ligation-independent cloning (LIC) system harbours an IPTG-(Isopropylthiogalactoside) inducable promoter, and it is based upon the treatment of a linearized vector backbone with T4 DNA polymerase in the presence of only one dNTP, (here ATP). The enzyme removes nucleotides from one strand of each specifically designed end until it encounters a residue corresponding to the dATP present in the reaction mix. The exonuclease activity of the enzyme will be overridden by its polymerase activity, resulting in specific non-complementary 13-14 base single stranded overhangs in the Ek/LIC vector. PCR products to be inserted are created by building appropriate 5' extensions into the gene-specific primers. The PCR product is further treated with T4 DNA polymerase, and then the unidirectional insert and the linearized vector are ligated through annealing. In this work, two recombinant prokaryotic pET expression vectors from Novagen were transformed into a bacterial BL21(DE3)pLysS strain, which lacks *Ion* and *ompT* proteases. The designation (DE3) indicates that the host is a lysogen of lDE3, and therefore carries a chromosomal copy of the T7 RNA polymerase gene under control of the *lacUV5* promoter. pLysS cells carry pET-compatible plasmids that encode T7 lysozyme, which is a natural inhibitor of T7 RNA polymerase. Cells containing pLysS produce a small amount of T7 lysozyme, used to suppress basal expression of T7 RNA polymerase prior to IPTG-assisted induction, and thus stabilize pET recombinants encoding target proteins that affect cell growth and viability.

Please see the appendix section for plasmid maps. Additional information regarding the LIC system and BL21(DE3)pLysS may be obtained at http://www.novagen.com. Fortunately, bacterial expression systems create huge amounts of the protein of interest (often 50% of the total bacterial protein content), finally taking control of the expression machinery and thus killing the cells. In our experiments, we therefore used spectrophotometric analysis to survey growth conditions, permitting us to induce COAS2-expression when cells were in exponential growth. We used Coomassie Brilliant Blue staining of 12% SDS-PA gels to visualize the results.

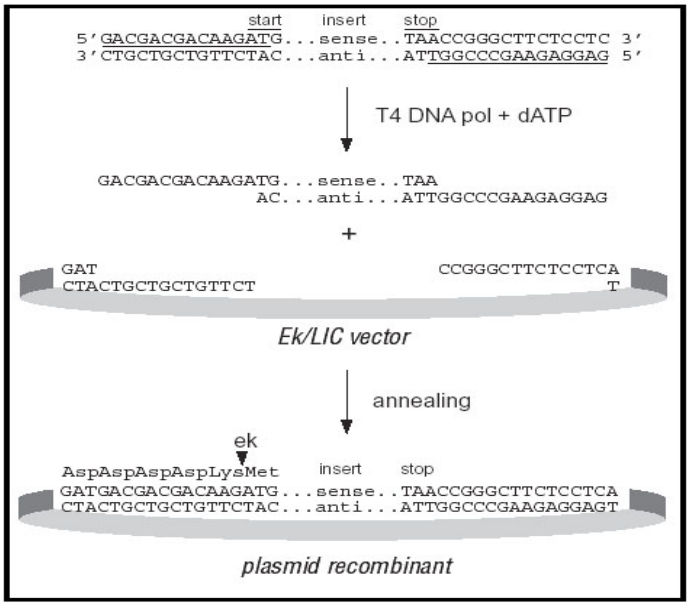


Figure 2.13: Illustration of the ligation independent cloning system, (Novagen). Ek = Enterokinase

Cloning procedure outline

pET30EkLIC-*COAS2* was cloned by Mr. M. Skårn and the author using LIC STARTCOAS2/LIC STOPCOAS2-primers in the Novagen Ek/LIC manual (#TB163 Rev A 0602) and transformed into BL21(DE3)pLysS as outlined in section 2.1.1. Colonies were PCR-screened as described in section 2.3.4. Mr. Jørn Henriksen made the pET3a-COAS2 construct according to the same LIC manual. After further verification by restriction cutting and electrophoresis, freezer stocks were made as outlined in section 2.1.1.

Growth and expression conditions in the bacterial systems

A tiny amount of freeze stock was picked with a sterile pipette tip, resuspended and grown in 5 ml LB for 12-18 hours at 37°C with 1% D-glucose, chloramphenicol (25 μg/ml final), kanamycin (25 μg/ml final) or ampicillin (50 μg/ml final). Kan and Amp added for the EkLIC30 and pET3a, respectively). The next morning, 1 ml of the culture was washed for 1 minute at 12000 rpm, and the pellet resuspended in fresh LB before it was transferred to an Erlenmeyer flask with 49 ml pre-heated LB with 0,25% D-glucose and Kan or Amp as earlier. Cells were grown at 37, 30 and 25°C in separate experiments, and 0,5 ml-aliquots were taken during growth. 0,1 ml was immediately harvested at 4000 x g for 20 minutes in previously weighed eppendorf tubes, and both supernatant and pellet were separately frozen in a mix of dry ice and 70% EtOH, and stored at -20°C. 0,4 ml was assayed for absorbance at 600 nm using the Shimadzu UV-1201 spectrophotometer with LB medium as reference. This made it possible to closely watch growth rates and to induce expression at a suitable cell density. At A_{600} = 0,6-0,7, IPTG was added. After 3-4 hours of induction, the reaction was halted. In some experiments, D-glucose was left out and addition of NaCl/$CaCl_2$ was attempted both with and without D-glucose to optimise the folding conditions for the expressed protein.

Preparation of bacterial lysates for protein SDS PAGE

The pellet's wet weight was measured before it was thawed on ice and resuspended in lysis buffer (see appendix for recipe) at 2-5 ml per gram wet weight. Lysozyme was added to 1 mg/ml, and the blend was kept on ice for 30 minutes before sonication (on ice) was performed at 25 W for 6 x 10 seconds with 10 seconds breaks in between. The viscous lysates were centrifuged at 10000 x g for 25 minutes at 4°C to pellet the cellular debris. The supernatant would now contain total protein. In the case of inclusion body formation, some lysates were submitted to a $Tween^{20}$/EGTA-cleansing cycle: A solution of 0,25 % Tween and 0,1 mM EGTA was mixed with the pellet suspension in question, followed by a 13000 rpm spin for 3 minutes at 4 °C. The supernatant was kept, while the pellet was washed twice more. Electrophoresis was performed as described earlier.

2.15.2 Protein expression in the RTS 100 cell free system

We also tried the RTS 100 cell free expression system from Roche Applied Science. Using both the pET3a*COAS2* and pET30Ek/LIC*COAS2* plasmids with the included protocol, we were able to obtain egfp control protein, but not to produce sufficient amounts of COAS2 for SDS PAGE visualisation with Coomassie Brilliant Blue staining. Please see the supplier's protocol for supplementary information.

2.15.3 Expression of *COAS2* in insect cells using the BAC-to-BAC system

In order to express our gene in an insect cell model system, the Bac-to-Bac system was chosen. This system relies on the generation of recombinant baculovirus by sit-specific transposition in *E. coli* rather than homologous recombination in insect cells (Jarvis et al., 1990);(Wickham et al., 1992). A new plasmid vector was created, inserting *COAS2* blunt ended without start codon into the *Ehe*I-digested FastBacHTa vector. DH10Bac bacterial cells were later transformed with this recombinant vector to yield a bacmid. After isolation, the Sf21 insect cells (*Spodoptera frugiperda)* were transfected with the bacmid. After 48 hours of incubation, recombinant baculovirus could be isolated. The viral stock would finally be used to transduce insect cells for protein expression, and protein lysates could ultimately be obtained. Please see www.invitrogen.com for further information regarding the Bac-to-Bac system. Dr. David J. Warren performed transformations, transfections, transductions and culturing of the cells. The author created the recombinant vector and performed the western blots required for protein detection.

Cloning procedures

An illustration of this procedure may be found in the appendix section. First, *COAS2* was amplified from the FRT/To/*COAS2* plasmid source using a 20-cycle PCR with Pfu proofreading polymerase and the COAS2UpperBac/LowerBac primers. The PCR product was gel purified according to section 2.2.4, and 2 μg of the gene-containing eluate submitted to end phosphorylation as described in section 2.3.5. The FastBacHTa vector had been digested with the enzyme Ehe1, leaving blunt ends. The plasmid was thereafter treated with shrimp alkaline phosphatase (SAP) to prevent re-ligation. Dr. David J. Warren of the Norwegian Radium Hospital performed the latter two preparations. The vector and insert were ligated in a

1:3 vector/insert ratio over night at 14°C as described in section 2.1.6, before the constructs were heat-shock transformed into DH5α cells. 20 colonies were screened for the COAS2 gene using PCR as described in section 2.3.4. COAS2Upper/LowerBac primers and a DNRTaq polymerase were used, and positive clones were grown in 5 ml LB over night with 50 μg/ml Ampicillin, and miniprepped as described in section 2.1.2. Having verified sequence and direction/correct reading frame, Dr. David J. Warren performed all further steps in the Bac-to-Bac system. Insect cell lysates were analysed using the western blotting method, see section 2.16.1. Attempts to enrich the cell lysates for HIS_6-COAS2 were made using immunoprecipitation (IP) with Protein G agarose (Sigma), see further on.

2.16 Expression of *COAS2* in mammalian Flp-In 293 cells

To enable localisation of COAS2 in the 293 cells, we aimed to use confocal microscopy (see sections 2.13 and 2.14 for more on CFM). A C-terminal Myc-His_6 tag was taken from the pcDNA3.1 B(+) Myc-His_6 expression plasmid (kindly provided by Hans Christian Åsheim, Department of Immunology at the Norwegian Radium Hospital). *COAS2* was PCR-amplified from the FRT/TO/*COAS2* vector with a DaveTaq polymerase and the UpperXbaI/LowerSfuI primers, followed by restriction cutting of both pcDNA3.1 vector and PCR product with the *Xba*I and *Sfu*I enzymes. Gel purification was followed by ligation as described in section 2.1.6, and again transformation into JM109 cells as described in section 2.1.1. PCR with the SfuI/XbaI primers was used for screening colonies, and stocks were ultimately made after restriction cutting and size verification plus automated dideoxy-sequencing. The next PCR step with XhoISTOP and BamHICOAS2 primers and DaveTaq polymerase gave an expected ≈ 600 bp product visualised with PAGE. After purification (section 2.3.3), the PCR product was digested along with the pcDNA5.1 FRT and FRT/To vectors with *Xho*I and *BamH*I enzymes (NEB) before gel purification and ligation over night at 16°C. Transformation into JM109 cells, PCR screening of the clones with specific primers, miniprepping, sequencing and maxiprepping was performed as described earlier. Each plasmid was co-transfected with the pOG44 plasmid into their respective 293 cell line in a 12:2 ratio of FuGene : DNA (μg). To verify the expression of our gene, we established a time series of tetracycline-induced samples. These were submitted to RNA purification with Trizol along with one constitutive 293 clone. All samples were further quantified, and northern filter was subsequently created. The filter was thus hybridised with a *COAS2* cDNA

probe. Further protein assays were performed as for the FastBac system, with the α-Myc 9E10 antibody in addition to the α-His antibodies.

2.16.1 Western blots and antibodies

Western blotting provides a direct method for identifying, monitoring, and determining the presence and relative amount of specific proteins in biological samples (Gershoni and Palade, 1982); (Renart and Sandoval, 1984). In comparison to other blotting techniques, (e.g. Southern and northern blotting), proteins are separated in a polyacrylamide gel, and blotted either by gravity forces or by the aid of electric force in a blotting apparatus onto either nitrocellulose or polyvinylidene fluoride (PVDF) membranes with a poresize of 0,45 μm. The filters are blocked towards non-specific hybridisation and further exposed to conjugated antibody solutions. Visualisation is possible through enhanced chemoluminescence (ECL) after thorough wash cycles. The wash solutions often contain extra detergents such as SDS or 'Tween^{20}' to minimize unspecific binding and thereby ameliorate the results. In a western blot, commercially available *antibodies* are used to detect native or recombinant proteins with high specificity.

2.16.2 Chemoluminescence

Emission of light from a substance in an excited state is defined as *chemoluminescence* if it is effected by a chemical reaction. The most known system may be the horseradish peroxidase (HRP)/hydrogen peroxide-catalyzed oxidation of luminol under alkaline conditions, which was used for visualisation of western blots in this work. The cyclic diacylhydrazide luminol is more efficiently oxidated by HRP in the presence of phenols, increasing the light output and emission time (also known as "enhanced chemoluminescence", or ECL). Light produced has a maximum emission at 428 nm, which can be detected by a short exposure to blue-light sensitive autoradiography film.

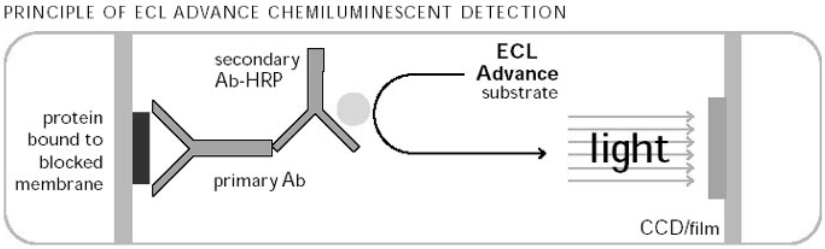


Figure 2.14: Illustration of the ECL advance system used for visualisation of Western Blots. Picture taken from the ECL kit protocol (Amersham Biosciences).

2.16.3 Preparation of mammalian cell lysates

The samples are most often non-purified total protein cell or tissue lysates. The denaturating lysate buffers contain cocktails of proteolytic inhibitors, detergents and chaotrophic salts to prevent proteolytic activity and hence fragmentation of the protein(s) of interest. It is important to keep all solutions on ice while working. In this work, both native (i.e. non-denaturating), and denaturating conditions have been used to lyse cells and retrieve their protein content. Electrophoresis, however, has only been performed under denaturating conditions (SDS PAGE with β-mercaptoethanol added to the sample loading buffers).

2.16.4 Protein lysate from mammalian cells in vitro

Cells were harvested as described in section 2.9.2, and the tube containing the pellet was snap frozen in liquid N_2 and stored at -70°C. Solutions for denaturating conditions were made according to guidelines in the Qiagen Expressionist, using 6M Urea as denaturating agent. Please see www.qiagen.com for additional information.

Method 1

The pellet was dissolved in lysisbuffer A[8], with added protease inhibitors (Pepstain A, Leupeptin Hemisulfate and Aprotinin at a final 0,01 mg/ml; PMSF to 1 mM). As a rule, 0,1 ml buffer was used per million cells. After 15-120 minutes on ice with occasional vortexing,

[8] 150 mM NaCl, 50 mM Tris-Cl pH 7.5, 0,1% NP-40

the samples were sonicated 3 x 10 seconds at 25 W and spun at 13000 rpm for 15 minutes at 4°C. The supernatant and pellets were frozen separately at -70°C.

Method 2

The cell flasks were deprived of medium and held on ice while washed once in ice cold PBS, before addition of 500 µl warm lysis solution[9]. Using a sterile rubber cell scrape, they were collected and further transferred by pipette to Eppendorf tubes. Boiled for 5 minutes at 100°C and again cooled on ice. The genomic DNA is sheared by mechanical rupture, resuspending the solution 4-5 times through a Microlane3 (0,5 x 16 mm) syringe tip. The solution is further centrifuged at 4°C for 7 minutes at 13000 rpm, and the supernatant is kept and an aliquote taken for total protein content analysis. After quantification, a protease inhibitor cocktail can eventually be added in a 1:100 dilution before the samples are snap frozen in liquid N_2 and transferred to -70°C.

2.16.5 Measurement of protein content (BRADFORD)

The total protein concentration was measured using the Bradford assay as first described by M. Bradford (1976). The standard microtiter plate protocol was followed, using a Wallac microplate reader at 595 nm along with bovine gamma immunoglobulin (IgG) as standard.

2.16.6 Dot Blot protocol

Serial dilutions of standards (and later samples) were used for antibody optimisation, diluted in lysis buffer and applied with a pipette onto dry pieces of nitrocellulose paper (BIO-RAD), and air-dried for 30 minutes. Further incubations and ECL visualisations proceeded as for tank blotted filters, see further on. We used primary antibody concentrations ranging from 0,1-1,0 µg/ml.

[9] 1%SDS, 1 mM sodium orthovanadate, (Na_3VO_4), 1 mM Tris-Cl pH 7.5

Protein SDS-PAGE

Our expected expression products had molecular weights of 18-25 kDa, and we chose 12% PA-gels for the SDS PAGE, which was performed as described in 2.2.3 with a dual color MW-standard from BIO-RAD. The 6x loading buffer contained 25% β-mercaptoethanol to prevent reformation of disulfide bridges. Samples were diluted to 20 µl in lysisbuffer, added loading buffer and denatured at 95°C for 5 minutes. Following a brief centrifugal spin to pellet debris, the samples were loaded onto the gel, which was cast as described in section 2.2.2 and 2.2.3, using the recipe from table 2.5. Electrophoresis was performed with a 1 x glycine/Tris western running buffer (appendix B), and the voltage was set to 30 mA for 15 minutes followed by 50 mA to end.

Western blot protocol (tank transfer)

The gels were transferred to a western blotting buffer (either Bjerrum-Scäfer-Nilsen, (BSN), or blotting buffer A, please see appendix for recipes), as the Mini-Protean^{II} blotting equipment (BIO-RAD) was assembled. When PVDF filters were used, they were soaked in MeOH for 1 minute, subsequently rinsed in doubly distilled water and 1 x blotting buffer. This was necessary due to its hydrophobic nature. Nitrocellulose filters were directly soaked in buffer before the blotting equipment was assembled as illustrated here.

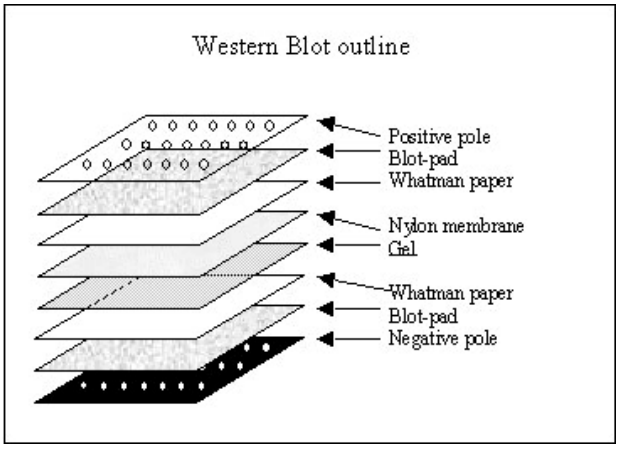


Figure 2.15: Illustration of a blotting apparatus for western blotting.

All blots were run at 4°C, with magnetic stir-circulation of the buffer and cooled with ice to keep the equipment from over-heating. The tank transfer was performed overnight at 30 V or 400mA for 60-70 minutes. Filters could be stained with Amidoblack for 5 minutes and destained with the Amidoblack wash solution for 2x10 minutes at this point. The filters were washed in ddH_2O to clear the inhibiting methanol. All blocking and incubations with different primary antibodies were performed according to the supplier's directives.

The following primary antibodies were used, all at 0,5 µg/ml final concentration: *In PBS/PBST with 0,5% dry milk: monoclonal mouse-α-human α-Myc 9E10 (kindly provided by Dr. David J. Warren), monoclonal mouse-α-human α-His_6 and monoclonal mouse-α-human α-HisHRP (Roche Applied Science). In TBS/TBST[10] with 3%BSA: monoclonal mouse-α-human α-pentaHis (Qiagen).* The Roche antibodies proved to be dysfunctional, not being able to bind positive controls under native conditions (results not shown). After a 3 x 10 min wash in either PBS/PBST, or TBS/TBST, the filter was exposed for 60-120 minutes to the secondary HRP-conjugated antibody from DAPI, diluted 1:5000 in PBS + 5% dry milk (w/v) or TBS + 5% dry milk (w/v). This was followed by 3 new 10 min PBST/TBST wash cycles, where the final wash was with distilled water only, before ECL-visualisation according to the supplier's protocol. The films were exposed from 10-600 seconds and developed as described earlier. The ECL kits were from Pierce (SuperSignal West Pico), and Amersham Biosciences. Solution recipes can be found in the appendix section B.

2.17 Immunoprecipitation of epitope-tagged COAS2

Immunoprecipitation (IP) is a procedure by which peptides or proteins that react specifically with an antibody are removed from solution and examined for quantity or physical characteristics (molecular weight, isoelectric point, et cetera) (Harlow, 1988); (Sambrook, 2001). Antibody-antigen complexes are removed from solution by addition of an insoluble form of an antibody binding protein such as Protein A, Protein G or second antibody. Analysis of the immunoprecipitate is usually by electrophoresis although other techniques can be used (Sambrook, 2001). The choice of immobilized antibody binding protein depends upon the species that the antibody was raised in. In this work, we applied Protein G, since it binds

strongly to IgG from cow, goat, sheep, cow, horse, rabbit and guinea pig and to mouse IgG_1 and IgG_3. Protein G can also bind bovine serum albumin (BSA). Thus, BSA was added to buffers used with Protein G. To detect the mouse IgG_1-antibodies used in this work, we used three different protocols from Qiagen and Invitrogen (www.qiagen.com, www.invitrogen.com). Experiments were conducted with HIS_6- tagged COAS2 from insect cells. These were initially detected using penta-HIS antibodies, which again was bound by the Protein G reagent. SDS-PAGE was applied for sample analysis. Please see the supplier's websites for further information regarding immunoprecipitation.

[10] T = 0,05% (v/v) Tween[20]

3 RESULTS

Cyclophilins have been extensively investigated over the past 20 years, resulting in findings that may be of value for the COAS2 protein. It has 84% amino acid homology to peptidyl-prolyl isomerase A (PPIA), a protein with diverse functions. PPIA has been reported to be important during developmental stages of many human malignancies (e.g. HIV-infection of host cells and cancer development), whilst *COAS2* has been found amplified and over-expressed in a subset of human cancers (Bukrinsky, 2002; Meza-Zepeda et al., 2002). To characterise this potential oncogene, we initially performed analysis of the nucleotide- and amino acid sequence, followed by diverse methods examining potential cancer-related effects of cells over-expressing COAS2 (e.g. cellular growth, cell cycle regulation and COAS2-mediated cellular resistance to chemotherapeutic agents). In parallel, we expressed the gene product in three model systems for production of a monoclonal antibody towards COAS2.

3.1 Sequence analysis of *COAS2*

It has been shown that PPIA harbours a proteoglycan-binding site at its C-terminus (Saphire et al., 1999). A Blast2 search verifies the existence of such a sequence in the translated *COAS2* mRNA, where the *148RNSKTSKK155* site differs in one amino acid only between the two: (150G→150S in *COAS2*). In addition to this, the pI of COAS2 has been estimated *in silico* to be 9.3, compared to the CypA pI of 7.8. Furthermore, a hydropathy plot of COAS2 versus that of Cyp A made using the Vector NTi software gives very different total hydropathy values (H_i-values) for these two proteins.

The search name "COAS2" in the NCBI Entrez protein database presently holds the accession number NP_839944 with REFSEQ-accession number NM_178230. Dr. Shinji Tanaka submitted an identical sequence to the NCBI in May 2002, in parallel with the *COAS2*-sequence submitted by Dr. Leonardo A. Meza-Zepeda. Dr. Tanaka reported the transcript as over-expressed in and thus related to liver cancer, and the sequence was termed "Cyclophilin-LC". Later findings at the DNA and protein levels have been limited. In 2003, Mr. Andrzej Galat summarised the findings of class A-cyclophilins; a subgroup to which *COAS2* belongs. In the human genome, the author reported the existence of 18 paralogues of the 18 kDa CypA *in silico*, whereof two have been identified and mapped to a chromosomal location (Galat, 2003). This number has increased to a total of 78 probable

http://pseudogene.org/human/index.html), whereof three locus-mapped pseudogenes, (named *PPIA1-3* (Willenbrink et al., 1995). The human CypA gene holds three alternatively spliced variants encoding two distinct isoforms that have been observed. A good number of the remaining paralogues could represent processed pseudogenes or mutated genes of some sort. However, most accession codes given in Galats paper have already been removed as a result of standard genome annotation processing, making further analysis difficult. In addition (until proven otherwise), an accession number does not necessarily represent a fully functioning gene that codes for a unique RNA or protein product(s).

To illustrate this point, potentially incorrect computer-based analysis indicates the existence of a long *COAS2* sequence found through computer-assisted genome-wide mapping and annotation ("GNOMON" at the NCBI). Only expression of the 495 nt mRNA has so far been confirmed using human tissue samples, even though GNOMON suggests that a larger mRNA of 217 aa with accession number XM_352323 should exist alongside *COAS2*.

The 660 bp RefSeq NM_178230 *COAS2* sequence (which harbours the 495 bp coding sequence), has a 75 bp leader sequence containing no significant repeated sequence plus one additional upstream start codon. The Kozak sequence of *COAS2* is 5' CTATCAGCC 3' The 1652 bp RefSeq NM_021130 PPIA sequence has a 72 bp leader sequence containing containing no significant repeated sequences nor additional upstream start codons, displaying a 5'CTAGTAGCC 3' Kozak sequence. The relevance of these leader sequences will be given further attention in the discussion module.

3.2 Artificial expression studies

The cyclophilin family of proteins exert diverse functions *in vivo*, and little could therefore be hypothesised about the novel family member COAS2. However, its high amplification levels in a subset of tumours indicated that it could be a novel proto-oncogene. Many proto-oncogenes are known regulators of growth (hereunder cell cycle regulation). Proto-oncogenes are thought to apoptosis and to cancer therapies. To examine effects of over-expressed COAS2 on characteristic cancer traits (growth regulation, chemotherapy resistance, cell cycle regulation and more), we decided to create stably transfected cell lines. Because FRT-

containing 293 cells were available, they were used for these studies although the cell type was not optimal for experiments related to mesenchymal oncogenesis.
In the egfp model system, the pEGFP-N1 plasmid was transfected into the 293 cells using the ratios depicted in table 2.6. Results indicated that electroporation under these conditions was less efficient than FuGene6. The results also showed that a 6:1 (µl/µg) ratio of FuGene6: DNA was the most efficient, with a maximum of 38,5 % living cells expressing egfp. This single experiment does not, however, eliminate experimental errors, and the relative percentages should therefore be merely viewed as estimates. Figure 3.1 shows the percentage of cells expressing egfp as a function of FuGene6:DNA ratio used for the transfection. The numbers were calculated using mean intensities for the fluorescent viable cells.

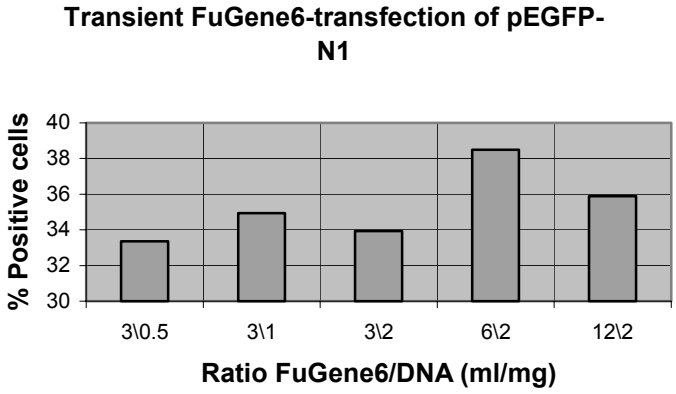


Figure 3.1: Percentage of living 293 Flp-In cells expressing pEGFP-N1 after transfection with FuGene6. Egfp-expression was screened by flow cytometry on transiently transfected cells.

The electroporation series, however, were not entirely adequate examples since the amounts of cells were $1/6^{th}$ of the recommended 4 x 10^6/ml in the standard protocols. For the two experiment series, we found that the optimal electroporator voltage setting was 54 V/mm, where 21% of the plated cells were successfully transfected (see figure 3.2).

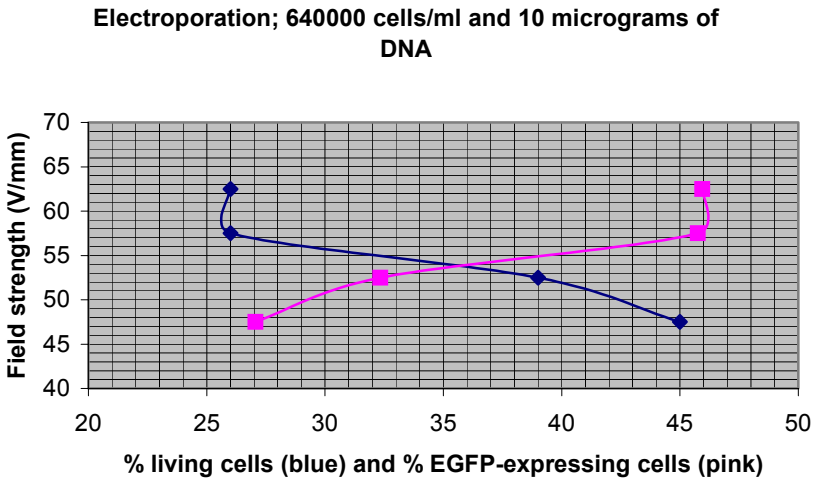


Figure 3.2: pEGFP-N1 electroporation experiment of 293 cells. The x-axis has dual values, acting as % of living cells for the leftmost curve, and as % of transfected (living) cells for the second dataset. See text for further explanation.

In a parallel experiment, parental cell lines in antibiotic-free growth medium were plated in 35 mm culture dishes for cytotoxicity assays, in a range from 1 x 10^5 to 3 x 10^5 cells per dish. six µl FuGene6 was added to 94 µl of serum free Dulbecco's Modified Eagle Medium (DMEM), followed by incubation for 15 minutes at room temperature and even dispersion of the blend into the growth medium of the plated cells. After 48 hours, the cells showed no apparent sign of growth inhibition or stress compared to the mock control (results not shown). Since the FuGene6 reagent caused no cytotoxic side effect, and demanded only an average of 2x10^5 cells per 35 mm dish for a single transfection (compared to 4 x 10^6 cells per ml in a standard electroporation experiment), we decided to keep working with FuGene6.

3.2.1 Testing the Tet-promoter system with EGFP-transfectants

Flow cytometry analysis of tetracycline-induced T-rex Flp-In 293 egfp transfectants was performed with parental T-Rex 293 cell line as negative control. The time points at 5, 24 and 96 hours were used to design the following curve (cells grown and harvested by Miss Lise Ramberg). The average intensity of green light was used, and its fold increase of intensity was calculated relative to the negative control. After tetracycline addition, the light intensity was found increased more than 1000-fold.

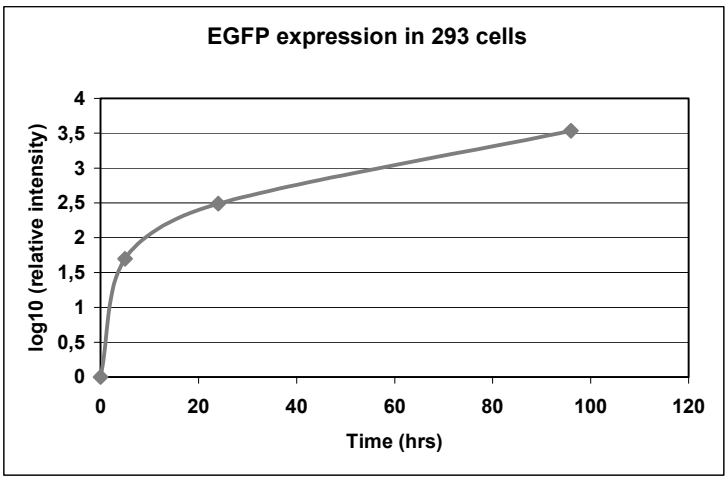


Figure 3.3: Increased intensity of green light upon tetracycline addition to 293 T-Rex cells transfected with egfp. Curve designed by Lise Ramberg.

The TetOn promoter did show signs of "leakage", as some of the uninduced egfp-transfected cells emitted green light when excited with blue light in an inverted fluorescence microscope. To investigate whether this originated from tetracycline present in the fetal calf serum (FCS) in use, a cell culture was grown in triplicate for 22 days in growth medium with 8% tetracycline-free calf serum from PAA. Flow cytometry was used for comparison of cells grown in the respective media. With standard FCS, we measured the amount of cells expressing egfp to 0,4% of the total. The new FCS merely lowered this value to 0,3% (results not shown), which may indicate that a permanent promoter leakage existed independently of serum tetracycline. We confirmed induction of egfp in the same clone by fluorescence microscopy (results not shown).

3.2.2 Creation of Flp-In cell lines and assay for antibiotic-sensitivity

In the Flp-In system, selection of positive clones is driven by addition of antibiotics to the growth media. Proper function and ideal working concentration of the antibiotic in use had thus to be evaluated before any transfection experiment could be launched. We therefore performed toxicity assays for Zeocin, Blasticidin and Hygromycin. One exception was made, where LOX-cells were chosen for Zeocin-testing, since our 293 parental cell lines were Zeocin-resistant. However, Zeocin does not have a traditional antibiotic "mode of operation" where cells round up and detach from the cell flask, but a slower progress where 293-cells

treated with Zeocin gain necrotic features with vague contours and blurred edges. The cells were therefore submitted to Zeocin-sensitivity testing with 1-1000 µg/ml Zeocin. After 2 weeks of selection, the apparently ideal concentrations were found to be 500 µg/ml Zeocin for sensitivity assays (five times the concentration recommended for culturing the parental cell lines), and 150 µg/ml Hygromycin for clone selection purposes. In a third experiment, both LOX-cells and 293 Flp-In parental cells grown in DMEM with 15 µg/ml Blasticidin were all dead after one week of treatment. This concentration was used to culture Flp-In T-Rex 293 cells and its descendants.

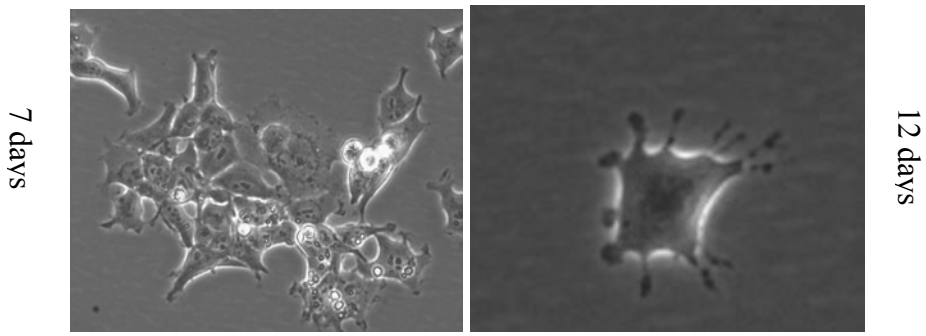


Figure 3.4: Zeocin killing 293 cells. Left: After 7 days of selection, right: single cell after 12 days of selection.

3.2.3 Stable transfection in the Flp-In system

FuGene6-transfection was used for the creation of all cell lines. After 24 hours of incubation, selection was initiated with medium containing Hygromycin (Flp-In 293 constitutive expression), and Hygromycin/Blasticidin (T-Rex 293 tetracycline-inducible expression). After 9-12 days, cell clones were picked and transferred using a 200-1000 µl pipette. Remaining cells were pooled after 17-20 days of selection. The clones were further mycoplasma-tested and assayed for Zeocin-sensitivity and finally stored as described in 2.9.4.

293 cells transfected with COAS2

Of 15 picked FRT/TO/*COAS2*-clones, only one was Zeocin-sensitive, compared to three out of the total eight for the constitutive clones. The reason why these clones showed resistance to

both Hygromycin and Zeocin was examined in a Southern blot analysis using DNA from parental cell lines and transfectants, see section 2.8.

293 cells transfected with *COAS2MycHis$_6$*

The two MYC-HIS$_6$-tagged *COAS2*-clones showed a higher degree of Zeocin-sensitivity. In the case of tetracycline-inducible clones, only the pooled backup samples collected in the final phase of cell clone isolation and expansion were resistant to Zeocin. The single expanded constitutive MYC-HIS$_6$ clone was Zeocin-resistant.

293 cells transfected with *egfp*

The Zeocin-sensitivity assay was not performed for the egfp-transfectants, although it may pinpoint a reason for promoter leakage (see further on). 12 clones were selected for each cell line; six with constitutive expression, and six with inducible expression. All clones were analyzed for egfp-expression using a Zeiss inverted fluorescence microscope with a FITC filter.

293 cells transfected with *COAS2/egfp and egfp/COAS2* fusion constructs

Mr. M. Skårn created two fusion constructs where the coding sequence of the *egfp* reporter gene was cloned N- and C-terminally of *COAS2* ORF, respectively. After cloning and purification of the recombinant plasmid, the sequences were verified. The constructs were subsequently transfected, followed by selection of positive clones according to section 2.9.6. Only two clones of 20 were found to be Zeocin-sensitive. All clones were further submitted to tetracycline induction (added to 0,1 µg/ml and 1µg/ml for comparison of egfp fluorescence intensity).

In parallel, Mr. M. Skårn transfected the egfp fusion constructs transiently into the Flp-In cells, i.e. without co-transfection of the Flp Recombinase-expressing pOG44 plasmid. As positive controls, the FRT/TO/egfp clone #1 (stable transfectants) and the pEGFP-N1 (transient transfection) plasmids were chosen.

3.3 Southern blot-assay of the Flp-In system

To assay the isogenicity and interclonal difference of the cell line, a Southern blot was utilised. The number of FRT integration sites in parental and clone could be connected to antibiotic resistance and sensitivity in the given clones. Since Zeocin-sensitive and Zeocin-resistant clones were found for both Flp-In systems, we therefore included both types of clones in this experiment, (constitutive and inducible expression). The hygromycin-gene of pcDNA5 plasmid-origin should produce a single band in all transfectant samples, and yield no band in the parental controls. The expected single band occurred in all samples for the *lacZ*-Zeo probe, indicating that both parental cell lines plus transfected downstream products had one copy only per locus of this gene. This was a prerequisite for cell line isogenicity. If the DNA was properly quantified but the band intensities showed signs of divergence, this could hypothetically be an indication that several copies had been integrated. The different intensities in the blots shown in figure 3.5 are nevertheless most likely due to concentration variations, (as shown in figure 3.6).

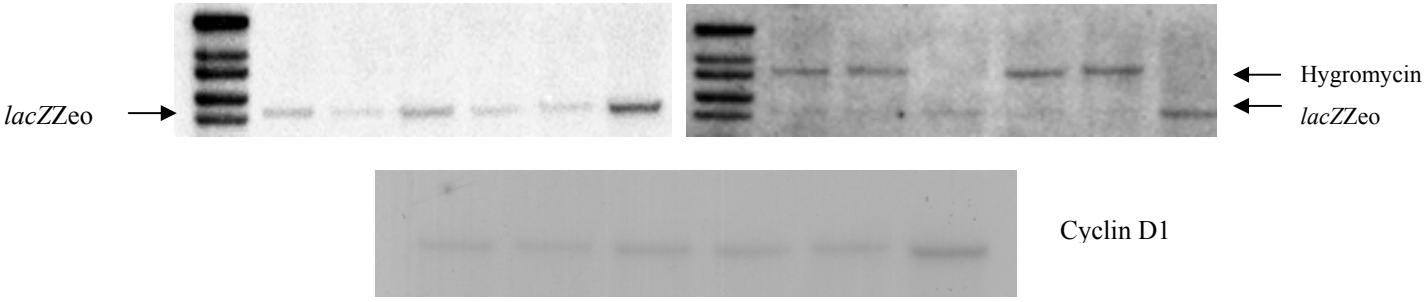


Figure 3.5: Southern blot of COAS2-transfected Flp-In 293 cells. Each lane harbours 7 µg *HindIII*-digested genomic DNA. Left: *lacZ*-Zeo-probe, and right: Hygromycin-probe. In both assays from left: 1kb MW-ladder, Constitutive #1, Constitutive #3, Flp-In parental, Inducible #12, Inducible #3, T-Rex parental. Some *lacZZeo* was left after on the filter even though the filter was stripped after the first hybridisation. This is indicated the figure (right). A Cyclin D1-probe was applied as loading control.

The blot was quantified using a fragment of the cyclin D1 gene, kindly provided by Ms. Åse Bratland. Laser densitometry was applied on the blot, as shown in figure 3.6.

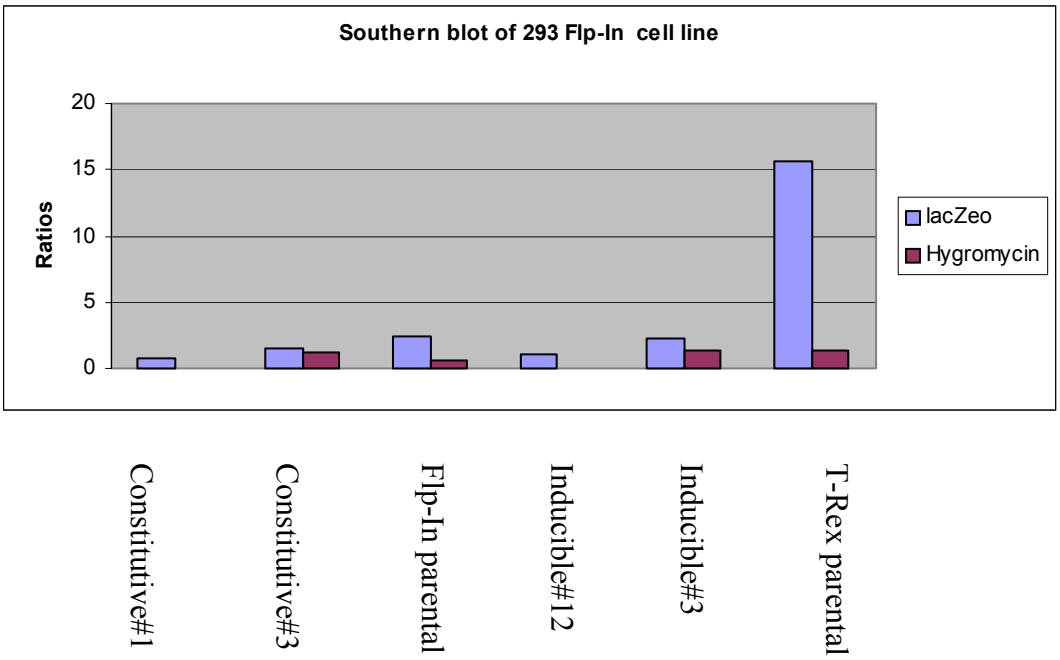


Figure 3.6: Quantified Southern blot from figure 3.4.

The levels of T-Rex parental cell line were very high compared to the other samples, explaining the strong band shown in figure 3.4 (*lac*Zeo probe).

3.4 Verification of *COAS2*-expression in the Flp-In system

To investigate whether *COAS2* was properly expressed upon tetracycline addition, a time series was established as described in sections 2.7 and 2.9. It should be considered, however, that tetracycline has a half-life of 24 hours, and must therefore be replenished every 48 hours to provide sufficient levels at all time points in the series. Six micrograms of total RNA was used for Northern analysis, along with equal amounts from the constitutive clone #1. The antibiotic resistance assays indicated that both the constitutive clone#1 and the inducable clone #12 were resistant to both antibiotics. The clone #3 was therefore used in all further experiments, except for the microarray experiments, where clones #3 and #12 were included to address potential dye incorporation differences.

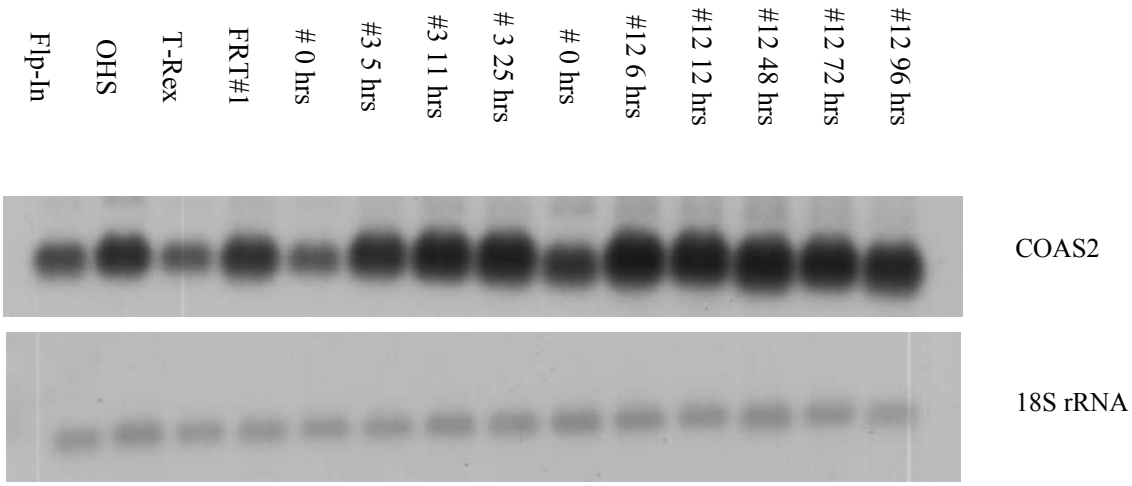


Figure 3.7: Northern blot of COAS2-transfected 293 cells. Top: tetracycline-induction of COAS2-transcription in tetracycline-inducable cell clones #3 and #12. Bottom: 18S r RNA loading control.

As indicated in the band at t = 0 hours, endogenously expressed cyclophilins may be present in the samples. Using the 18s rRNA and a laser densitometer, the northern blot was quantified: (Figure 3.8)

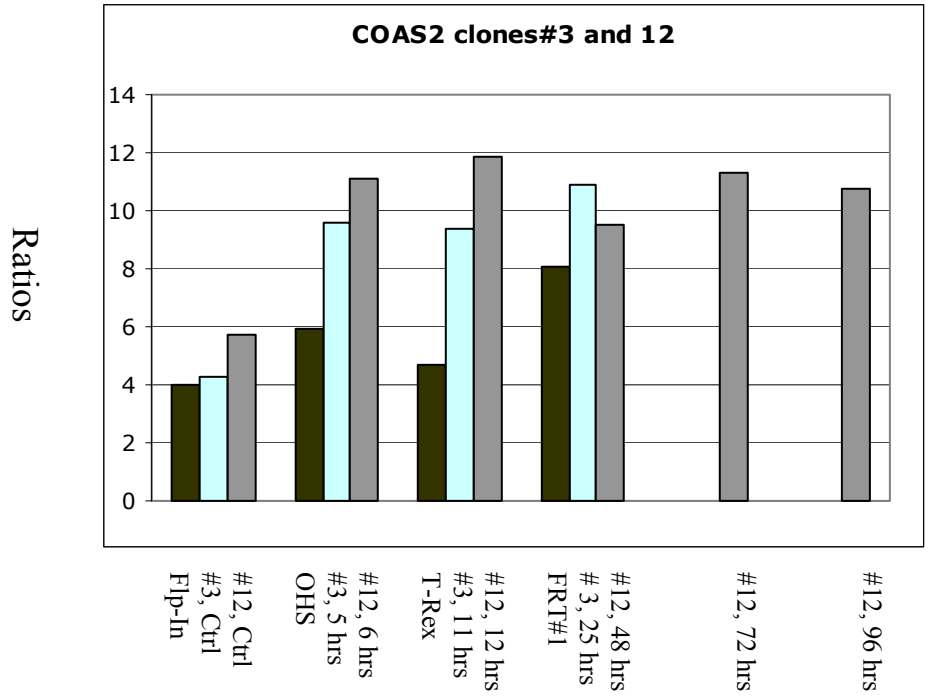


Figure 3.8: The expression levels of COAS2 are slightly clone-dependent. Tetracycline has a half-life of 24 hours, and was only added anew after 48 hours. Black: Parental cells and OHS, a cell line over-expressing *COAS2*. Grey: Time series of clone #12 with over-expressed *COAS2*. Turquoise: Time series of clone #3 with over-expressed *COAS2*.

3.5 Expression of *COAS2- MYC9E10-HIS$_6$* in the Flp-In system

COAS2, C-terminally tagged with MYC-HIS$_6$ was cloned into two different pcDNA5-plasmids, and further transfected with FuGene6 into the Flp-In cell lines as described earlier. A time series was made for northern blot verification of proper system function. RNA was prepared and the northern run as outlined in the methods section. The cDNA probe used for the *COAS2*-cell line material was applied to a filter containing ten micrograms of total RNA.

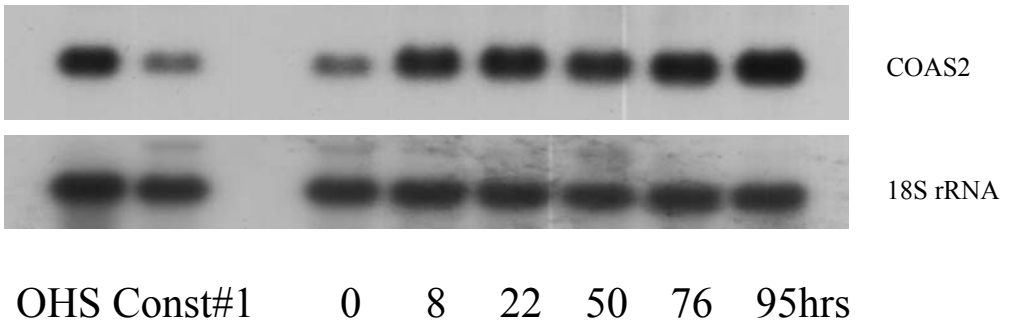


Figure 3.9: Northern blot illustrating the expression of *COAS2-MYC-HIS$_6$* mRNA in the tetracycline-induced samples compared to the negative control. The Const#1 is the clone with constitutive expression of COAS2. OHS is a cell line established at the Department of Tumor Biology from a primary osteosarcoma from femur. Equal levels of expression in both the Const#1 and the negative control sample (0 hours) cannot be explained without further experiments, please see further on. Fresh growth medium with tetracycline was added after 48 hours. Numbers from 0-95 signifies timepoint of sample harvest (hours after tetracycline addition). See figure 3.9 for additional details.

As indicated in the band at t = 0 hours, endogenously expressed cyclophilins may be present in the samples. To determine what the band represents, we must hybridise the filter with a probe containing the Myc9E10-His$_6$-sequence only. If a negative control, e.g. a parental 293 RNA sample is included; hybridisation of the Myc9E10-His$_6$ probe will probably determine the wild type identity of these bands. The low levels of expression in the Const#1-sample will also be addressed through a second experiment. The filter was hybridised with a control probe (18s rRNA), and quantified using a laser densitometer (Figure 3.9).

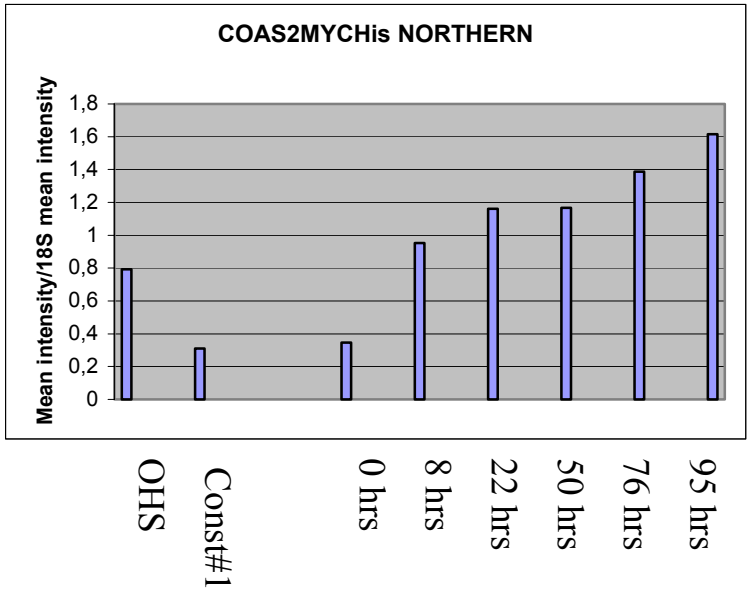


Figure 3.10: Diagram depicting expression ratios when quantified with a 18S rRNA control probe. Const#1 = cell clone with constitutive expression of COAS2. OHS is a cell line with amplification of *COAS2*. See figure 3.8 for additional details.

Although the existence of the Myc-His_6 tail was only verified for the plasmid DNA sequence, the northern filter strongly indicated a positive induction of our ORF when hybridised with a *COAS2*-probe. (We must hybridise the filter with a probe containing the Myc9E10-His_6-sequence only to address the existence of the tag.) We therefore proceeded to the protein visualisation step without further delay. Total protein lysates were prepared as outlined in the methods section, giving us two sample sets that were examined for COAS2-MYC-HIS_6 expression using SDS-PAGE. We never detected the protein, despite countless efforts of optimisation, using phosphate-and Tris-buffered saline of various concentrations, different antibodies and antisera, plus two different types of ECL-kits. For both α-Myc9E10 and α-His_5 antibodies, the positive controls gave strong, clear signals (results not shown). Possible explanations are hypothesized in the discussion section.

3.6 Growth rate and cytotoxicity experiments

3.6.1 Monolayer growth assays

In addition to the soft agar experiments, two *in vitro* growth rate experiments were performed using SulphoRhodamine B staining in order to calculate an approximate doubling time for

two COAS2-transfected cell lines (the constitutive clone#1 and the tetracycline-inducable clone#3), and the parental Flp-In 293. The first experiment would give an estimate of cell growth, and thus the ideal range of dilutions to plate. Data was hence obtained in the second experiment. Using data from the 3000 cells per well-series and the formula depicted in section 2.10, we were able to estimate as follows:

Table 3.1: Determination of cell doubling time in the *COAS2*-transfected clones using the SRB colorimetric assay.

Cell line	Abs_{540} (t = 0)	Abs_{540} (t = 4)	Doubling time
Flp-In 293	0,145	0,764	40.0 hours
Constitutive clone#1	0,175	1,044	37.25 hours
Tetracyline-induced clone#3	0,132	0,742	38.50 hours
Clone#3, negative control	0,106	0,576	39.3 hours

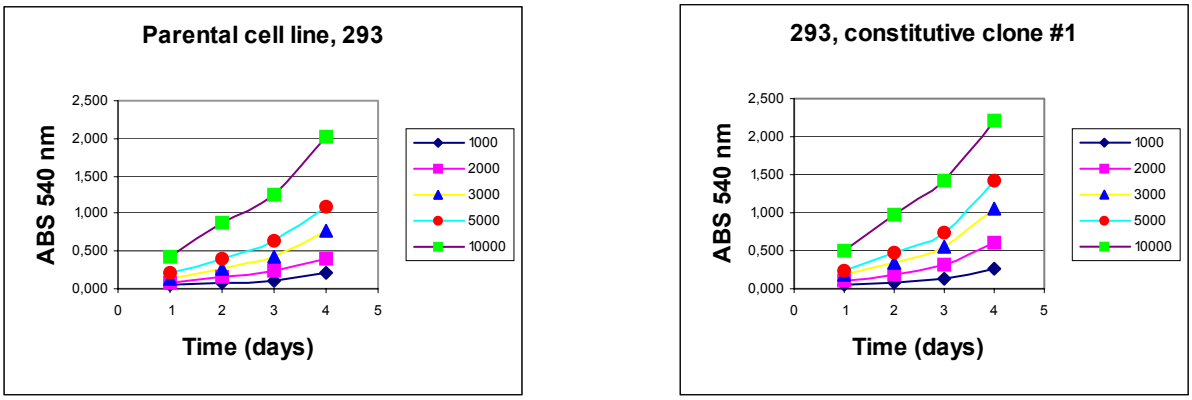


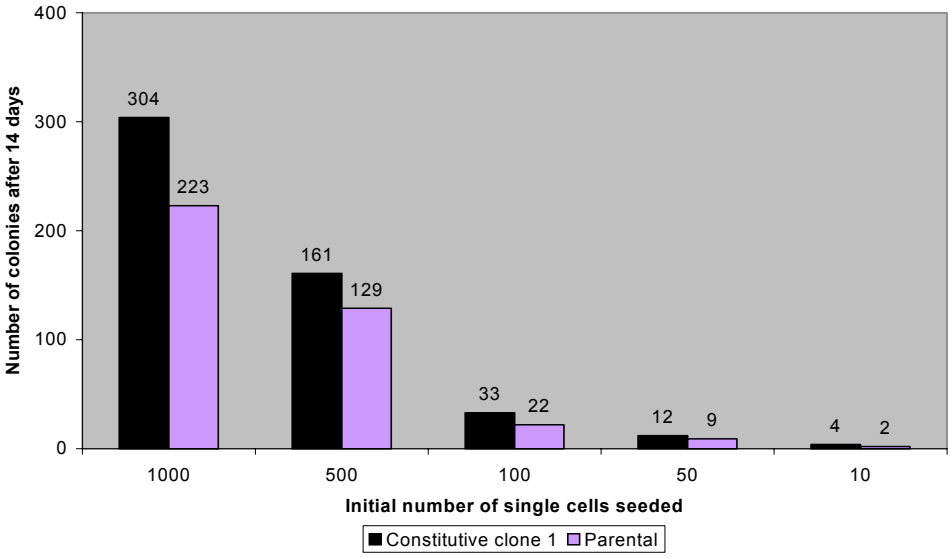
Figure 3.11: Growth curves for *COAS2*-transfected clone and parental cell line, respectively. The mid-range of 3000 cells per well was used to estimate doubling times.

Determination of growth rate for the 293 transfectants indicated that COAS2 had a small effect on growth rate when compared to a negative control.

3.6.2 Soft agar growth assays

With kind assistance from Mr. Meng Yu Wang, an initial experiment compared soft agar growth of the constitutive COAS2-clone #1 and the parental Flp-In 293 cell line. The first soft

agar experiment indicated a difference in clonogeneicity; the ability to form tumour-like cell colonies from single cells when grown without anchorage (attachment to a surface). The second experiment confirmed the first, even though minor internal differences for the constitutive clone data indicate that the number of initial cells differed in the two separate experiments. The data from the second experiment is shown, where triplicate mean values have been applied for each sample



.

Figure 3.12: Soft agar assay of cells over-expressing COAS2. The clone with constitutive expression forms more cell clones than its parental cell line, as indicated. Black = constitutive clone 1, and purple = 293 Flp-In parental cell line. Shown is one of two parallel experiments. Numbers are means of triplicate values.

3.6.3 In vivo growth assays

However, it is quite possible that COAS2 affects tumour growth *in vivo*, as indicated by the soft agar assay and monolayer results. We thus generated tumours from four 293 cell line transplants, by injecting cell suspensions subcutaneously in immunodeficient mice (athymic Balb-c/(nu/nu), bred at the Norwegian Radium Hospital). These xenografts were later propagated and their growth rates assayed over a time period of three months. Mr. Christoph Müller generated independent growth curves. The experiments were performed in triplicate, according to a protocol by (Kistner et al., 1996).

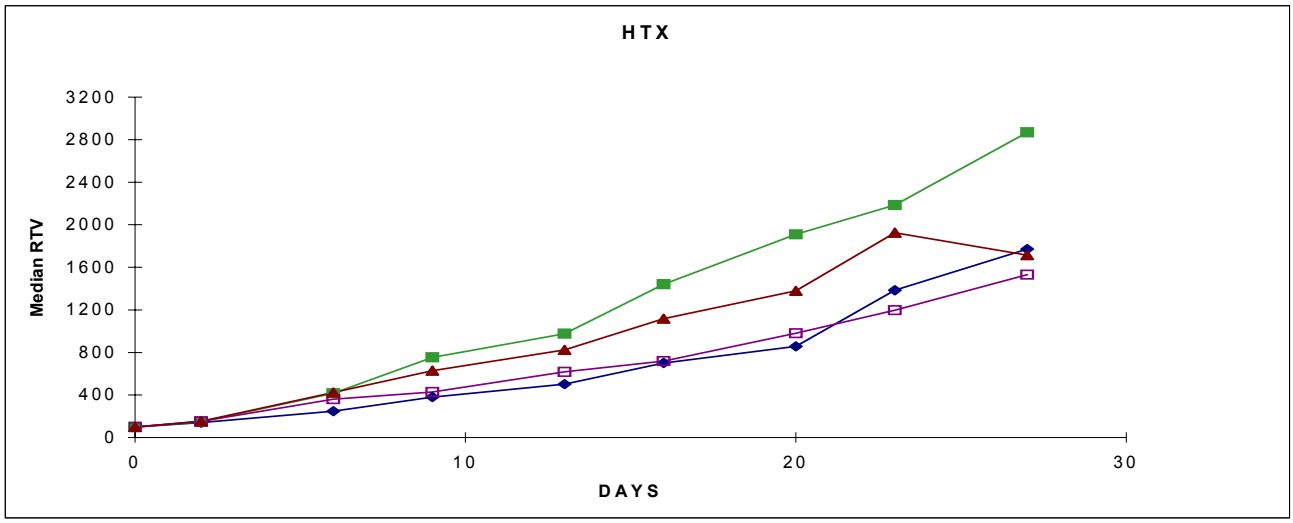


Figure 3.13: Model depicting growth rates for 293 cell grafts in nude immunodeficient mice. The tetracycline-inducable cell line FRT/To/COAS2/3 in green □, with the control tumours shown in red ∇. The constitiutive clone FRT/COAS2/1 (pink □), and the parental Flp-In 293 control (from passage 2) in blue □.

The growth curves of the xenografts indicate that over-expressed *COAS2* has a potential effect on tumour growth. In the tetracycline-inducible tumours, the Flp-In system harbouring the TetOn promoter may have been disturbed after a long period without antibiotic selection. Interestingly, the experiments show that both negative control and tetracycline-induced cells from the T-Rex 293 origin grow faster than both parental 293 and clone with constitutive expression of *COAS2*.

Total RNA from the tumours will be isolated and compared to cell line material using micro array analysis in the near future. Unfortunately, this was not possible within the timeframe of this work.

3.6.4 Cytostatica-sensitivity assays

A newly discovered oncogene may be of great worth, as its gene product may be sensitive to existing and new cancer therapies. To assay whether COAS2 could confer resistance to cancer drug treatment, Mr. C. Müller performed cytotoxicity experiments in triplicate with the stably COAS2-transfected clone#1 (constitutive expression), and reference parental 293 cell line. All samples were incubated with cytostatica for either four hours or 72 hours in triplicate experiments. The chosen cancer drugs cisplatin, doxorubicin (0,01-1000 nm) and

methothrexate (0,01-100 mM) gave similar results in this system, as sample and reference responded almost identically in all three experiments. Data from these experiments were summarised by Mr. C. Müller, and kindly reprinted here.

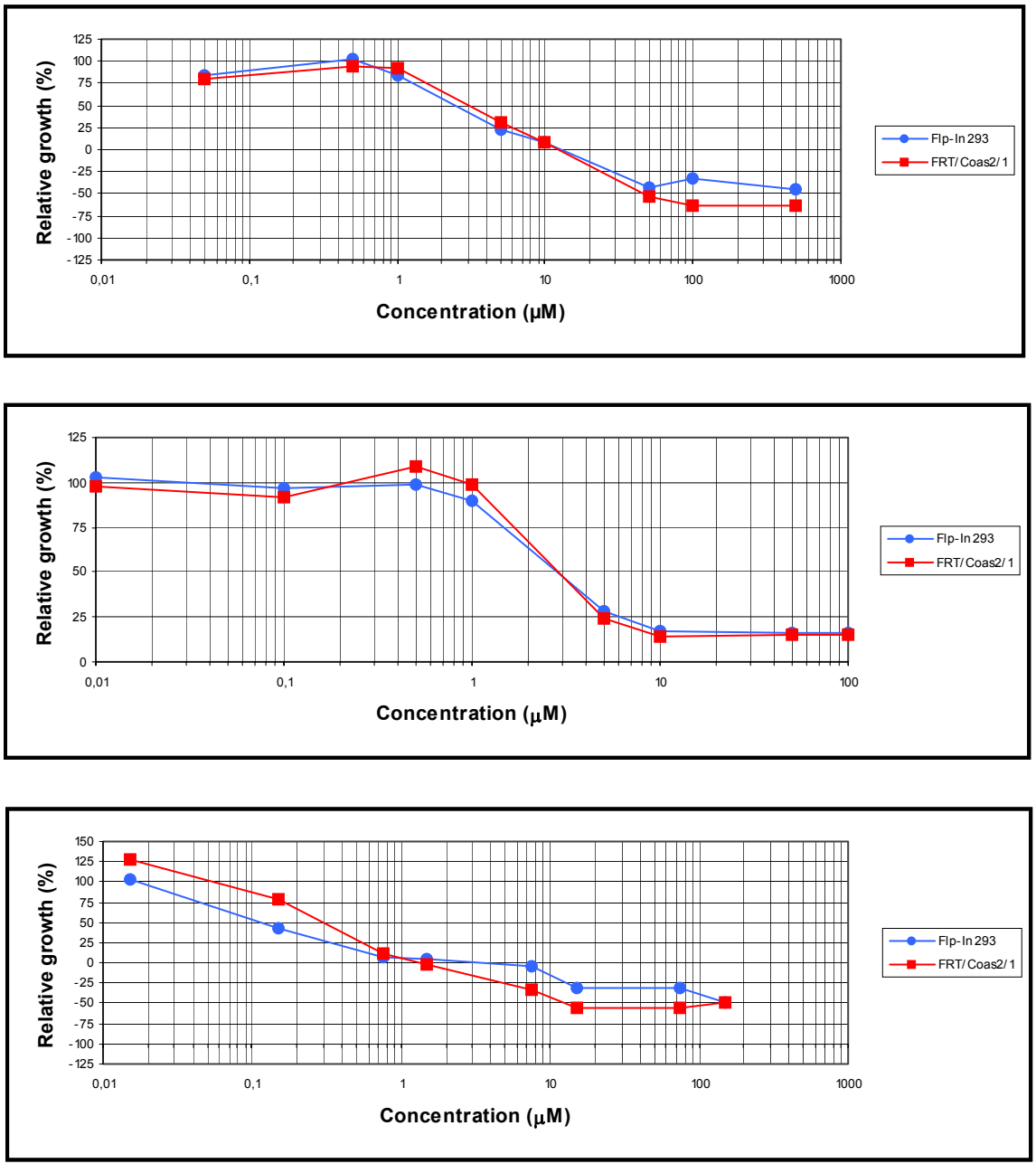


Figure 3.14: SRB growth inhibition assays. From top: Parental Flp-In 293 versus clone #1 with constitutive COAS2-expression, treated with cisplatin, methothrexate and doxorubicin, respectively. There were no significant growth differences between the cell lines in either experiment, indicating that over-expression of COAS2 in these cells may not confer resistance to these anticancer drugs upon treatment. Graphs and data are courtesy of Mr. C. Müller.

In all three assays, the relative growth rates were equally affected for parental and clone, respectively. This indicates that the over-expressed COAS2 gene product does not contribute to resistance to any of these drugs, at least not in 293 cells. The levels of COAS2 in these samples before and after treatment should ideally be verified in a repeated experiment.

3.7 Flow cytometry analysis of cell cycle ratios in *COAS2*-transfected cells

To investigate wether over expressed COAS2 had an effect on the cell cycle, we performed flow cytometry on methanol-fixed COAS2-transfected cells stained with the DNA-binding dye propidiumiodide. The 293 cells were prepared as described in section 2.13, with triplicates for all sample types. The experiment was repeated twice, and the graphs were visually compared. Thus quantitative data for the G1, M and G2 phases are not available, but will be upon repetition of the experiment using lower amounts of FCS in the cell growth medium. An interesting observation was made in the experiment using 2 % FCS; the existence of a small peak prior to the G1-peak in the negative controls only (indicated with an ellipsoid in figure 3.18 D). This could be background noise, or dead (possibly through apoptotis) cells stained with propidiumiodide. Please see discussion for further comments on this subject.

The cell lines fed with 2 % FCS grew more slowly than those with 8 % FCS, as expected. However, the 2 % did not fully inhibit cellular growth, nor did the samples grown in 2% FCS show any major difference to the 8 %-samples regarding possible COAS2-influence on the cell cycle display. (The tetracycline-induced cells showed similar G1-S-G2-patterns as the reference cells.) Although 2 % FCS caused lowered growth, an actual effect of COAS2 on cell cycle distribution may not be seen unless the experiment is repeated with 0,5-1 % FCS for more pronounced growth reduction. This was however not possible within the timeframe of this work.

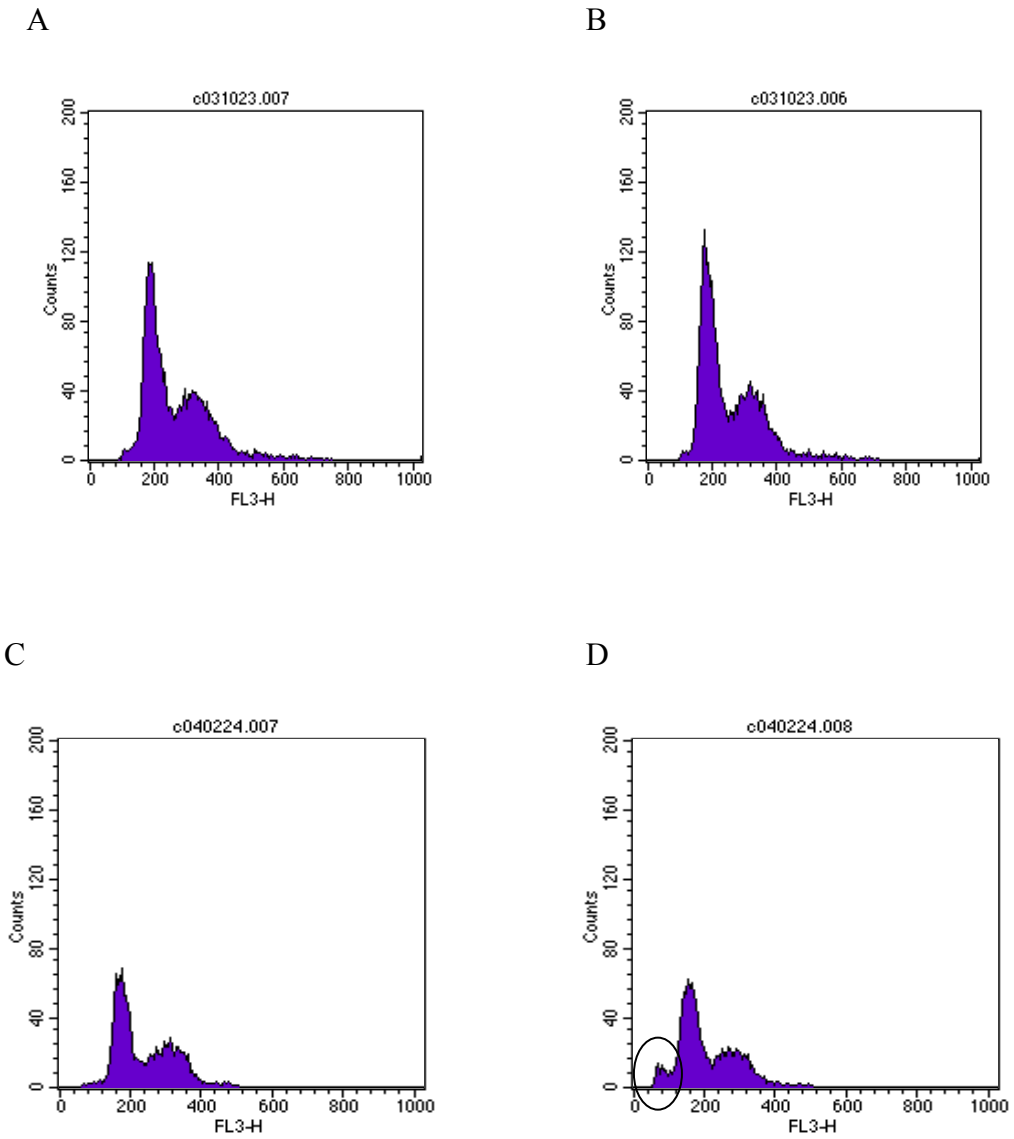


Figure 3.15: Cell cycle distributions of 293 Flp-In cells over-expressing COAS2. Y-axis: 5000 counted cells, x-axis: total fluorescence intensity (propidiumiodide excitation). Data from two unique experiments are shown. A: Standard conditions (8 % FCS), sample with over-expressed COAS2. B: Standard conditions, uninduced control. C: Cells fed with 2 % FCS, sample with over-expressed COAS2, and D: Cells fed 2 % FCS, uninduced control.

3.8 cDNA microarray-results

cDNA microarray analysis of stably transfected cells was applied in this work to create a transcription profile of cells over-expressing COAS2. Upon over-expression in 293 cells, the microarray procedure could potentially reveal signal pathways affected by ectopically expressed *COAS2,* giving information about the functional role of this cyclophilin.

Two series of cDNA slide prints were applied to the total RNA in question, and hybridisations have been performed at various temperatures, as manual and machine-assisted hybridisations. Initially, cDNA arrays were chosen for two different time series with three timepoints each (5,11 and 25 hours, respectively). This would give a notion of biological variation, as the series originated from two different clones. As reference in these series, non-induced samples harvested at the corresponding time points were used. Existence of *COAS2* mRNA had been verified in both clones, but one was resistant to both Hygromycin and Zeocin after isolation, whilst the other was only Hygromycin-resistant (#12 and #3, respectively, see figure 3.10). The RNA quality of the 25 hour sample from clone #12 was however not adequate, forcibly reducing the number of time points to five. In addition to the time series, the parental cell line was run as reference when analysing clone#1 (with constitutive COAS2-expression), an assay that theoretically should give the two series an endpoint for comparison (where a high level of COAS2 had been expressed for a long time). To determine the efficiency of dye incorporation and eventual downstream effects thereto, a single dye swap experiment was used for the sample set of constitutive expression. The dyes were equally incorporated in both sample and reference, hence not affecting the expression ratios for most transcripts. However, full accuracy only can be assessed after internal array-array normalisation, which was not accomplished in BASE.

The results are presented as sample/reference ratios of a subset of the upregulated genes, as well as total intensity values. *COAS2* was not present in the cDNA arrays. All the sequences of genes included (except the cyclophilins) have been verified.

Table 3.2: Ratios from cDNA microarray time series experiments. C = constitutive expression of COAS2. Ratio cutoff = 2; X = absent (flagged spot), L = spot with expression ratio of less than 2 in the experiment. Cl = cell line. Ratios stem from parallel or triplicate experiments except for the 6 and 12 hour samples.

Gene symbol	Gene name	Sample/reference ratio (Cy5/Cy3)					
		Cl: #3	#12	#3	#12	#3	
		5	6	11	12	25	C
COL1A2	Collagen type 1 alpha 2	L	4.4	L	4.6	L	2.2
GJA5	Gap junction protein, alpha 5, (connexin 40)	4.5	4.2	L	3.0	2.2	4.1
HMP19	Hypothalamus golgi apparatus expressed 19 kDa protein	4.1	4.6	L	3.1	L	3.4
IL7R	Interleukin 7 receptor	3.1	3.0	2.3	4.0	2.1	2.8
L-Sox5	SRY-box 5,transcript variant 2	2.6	2.9	3.0	5.9	3.8	4.1
MAP4K1	Mitogen-activated protein kinase kinase kinase kinase 1	2.9	3.0	X	3.4	3.5	X
MGLL	Monoglyceride lipase	X	4.7	2.6	4.9	2.1	L
MST1R	Macrophage stimulating 1 receptor	X	X	L	4.7	3.6	2.5
MYL3	Myosin, light polypeptide 3	2.1	3.9	L	3.2	L	4.5
Sox6	SRY-box 6	3.6	5.2	2.1	6.2	X	3.7
PPIC	Peptidyl-prolyl isomerase C	1.0	1.0	0.6	1.1	1.2	1.0
PPIF	Peptidyl-prolyl isomerase F	1.1	0.8	1.0	0.9	1.1	1.1
PPIA	Peptidyl-prolyl isomerase A	1.0	0.9	1.4	1.3	1.1	1.4

The 6 and 12 hour time points stem from the clone #12, whereas the three other time points originate from clone#3. The data for L-Sox5 illustrates a problem that may be caused by this fact: In the 12 hour-sample, the ratio value is twice that of the 11 hour sample, suggesting considerable clonal variation. However, northern blotting did demonstrate that levels of *COAS2* expression in clone #12 are higher than those of clone#3. The high level of COAS2 in clone#12 could contribute to the higher levels of several putative downstream genes observed here.

Table 3.3: Intensities taken from two separate experiments for the L-Sox5 and Sox6 genes. Background levels have been subtracted; intensities stem from normalised data.

Gene symbol	Gene name	Intensity 11 hour sample	Intensity 11 hour control	Ratio	Intensity 12 hour sample	Intensity 12 hour control	Ratio
Sox6	SRY-box 6	1521	504	3,0	3966	668	5,9
L-Sox5	SRY-box 5, transcript var. 2	805	383	2,1	2890	464	6,2

The interclonal variation were not adjusted for, (that is normalised array by array), normalising parallels and excluding spots that do not exist in 80% of the data set. Such a method is not available in BASE, but in external programs, creating a file that cannot be reverted for later data comparison in BASE. Importantly, most of the spots are ignored because of similar levels throughout the experiment. It should be mentioned that in-house printed cDNA arrays have been reported to have some errors. The most interesting clones in this context were sequenced (but not *COAS2*), and certain errors were found. The cause of these faults may have been in the clone collections or arisen during probe preparation.

In most assays, myosin light chain 1 (MYL1), interleukin 6 receptor (IL6R), tumour necrosis factor ligand superfamily member 11 (TNFSF11), and antigen CD37 were present as moderately upregulated genes. These probes were however not sequenced. The MGLL and HMP19 protein products were not examined further in this work. The collagen type one alpha two (COL1A2), was not present in more than half of the samples, but was included here because of its connection to the Sox family of transcription factors.
The cDNA arrays indicated two HMG-box-containing transcription factors L-SOX5 and SOX6 as the most highly upregulated in many of the samples. (Sequencing identified the L-*Sox5* and *Sox6*-probes to be of murine origin, added by a user. *M. musculus Sox6* and *L-Sox5* have 89 and 93% base pair homology to *H. sapiens* genes, respectively.) These two group D Sox proteins have been shown to form homo-or heterodimers through their coiled-coil domains and recognize pairs of HMG-binding sites (Lefebvre et al., 1998). Together with Sox9, the L-Sox5 and Sox6 proteins have been reported to regulate the expression of collagen type two alpha one (COL2A1) during chondrogenesis (Lefebvre et al., 1998). The COL2A1 was printed on the microarrays, but was not found to be significantly upregulated.

Expression of L-*Sox5*, *Sox6* and COL2A1 in our samples was examined using the northern blotting method. Clones for *H. sapiens L-SOX5*, -6 and -9 were purchased from the Deutsches Ressourcenzentrum für Genomforschung (RZPD) in Germany, and the corresponding *SOX5*, *SOX6* and *SOX9*-probes were generated as described in section 2.6.4. The *COL2A1* probe had earlier been shown to have sufficient sensitivity in Southern blots, and was kindly provided by Dr. Leonardo Meza-Zepeda. Through database searches, mRNA levels in mammals of the respective *SOX* genes were found to be low, in the range of 5-20 transcripts per million in a mammalian cell, whilst the collagen is only expressed at high levels during early stages of chondrogenesis. No bands could actually be seen when hybridised towards a RNA filter

containing Flp-In cell line material with over-expressed *COAS2* (Filter shown in figure 3.6).

The absence of bands could potentially be due to poor incorporation of α-^{32}P dCTP; therefore, the sensitivity of the *SOX*-probes were examined in a Southern blot: All three SOX-probes were hybridised to a filter containing genomic DNA from *H. sapiens* placenta and four cell lines whereof two were of human and two of murine origin. When the Southern filter was hybridised with *SOX6*, a single band occurred for all samples (results not shown), with the exception of two bands for the BALB-c murine cell line. L-*SOX5* produced signals for murine samples only, whereas *SOX9* did not hybridise at all to the filter. However, since the *L-SOX5*-probe only detected murine DNA when hybridised to a Southern filter, this could mean, of course, that the probe actually was from *M. musculus*, rather than *H sapiens*. It might be explained by human error during probe preparation either at the site of purchase (Ressourcenzentrum für Genomforschung (RZPD), Germany), or by the author. At the time, the conclusion was however that the SOX probes were difficultly labelled with γ-^{32}P. Nevertheless, (with only the *COL2A1* and *SOX6*-probes found to be adequately sensitive), we were not able to verify cDNA microarray results using northern blotting.

3.9 Intracellular localisation experiments using egfp fusion proteins

Mr. M. Skårn transiently transfected U2OS cells with pcDNA5FRT/TO/egfp/COAS2 and pcDNA5FRT/TO/COAS2/egfp. Both experiments displayed COAS2 fusion proteins (figure 3.16). The fluorescence was relatively weak compared to the pEGFP-N1 control, but it was similar for both fusion constructs. Located to the cell membrane and the cytoplasma, the speckled pattern obtained in earlier experiments by Ms. Lygren was not reproduced.

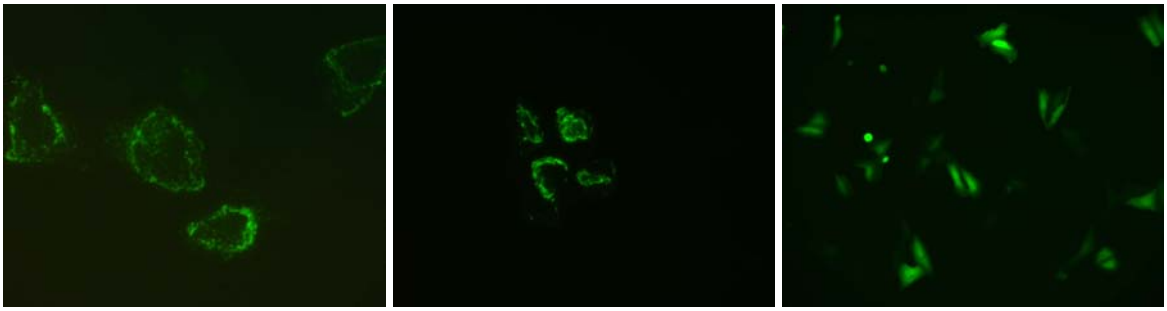


Figure 3.16: U2OS cells transiently transfected with COAS2/egfp and egfp/COAS2. Pictures were taken 17 hours after transfection. Left: COAS2/egfp, middle: egfp/COAS2, and right: pEGFP-N1 control. Pictures courtesy of Mr. M. Skårn.

However, when the same constructs were used to create stably transfected Flp-In 293 T-Rex cells, the cells showed no expression of fusion construct (no visible fluorescence above the autofluorescence from the 293 cells). Analysis of the stable transfectants using northern and western blotting will be performed, but it is outside the timeframe of this work.

3.9.1 Intracellular localisation experiments using COAS2-MYC-HIS_6

Intracellular identification of COAS2-MYC-HIS_6 was attempted using confocal microscopy on fixed cells stained with a Rhodamine-tagged secondary antibody, in addition to two separate experiments that addressed filamin and actin expression in fixed cells with over-expressed COAS2. The cells were prepared according to the standard protocol outlined in section 2.14.1, using in total three different primary antibodies, towards MYC, His_6 and filamin(s).

With the His_5- and Myc9E10 antibodies, attempts were made to localise the tagged COAS2 protein. The antibodies were initially applied in a range from 0,5-5 µg/ml, with the secondary Rhodamine-tagged antibody in a 1:100 dilution. Only the HIS_5 antibody produced a signal. This signal was localised to the nucleus (excluding the nucleolus), as visualized by adding 3 µl Vectashield mounting medium with DAPI-stain of the same cells. (DAPI stains genomic DNA and emits blue light in contrast to the red-emitting Rhodamine.) However, the signal arose in control 293 Flp-In parental cells too, indicating that this was a background artefact. The MYC9E10-signal was very weak, (at levels equal to control cells, whose levels were those of background signal only), and uniformly spread. Of course, endogenous MYC should

ideally have been detected if expressed in the 293 cells. The concentrations of primary antibodies were altered, using 5-10 μg/ml with secondary antibody concentrations as in earlier experiments. However, this did not produce a significant change: we were not able to detect the tagged protein.

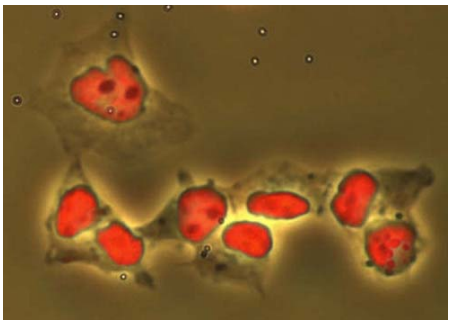


Figure 3.17: Cells with over-expression of COAS2-MYC-HIS_6 show a nucleus-specific (excluding the nucleolus) signal when stained with α-His_5 antibody. Secondary antibody was Rhodamine-tagged. Depicted is a merged photo taken after 16 hours of induction.

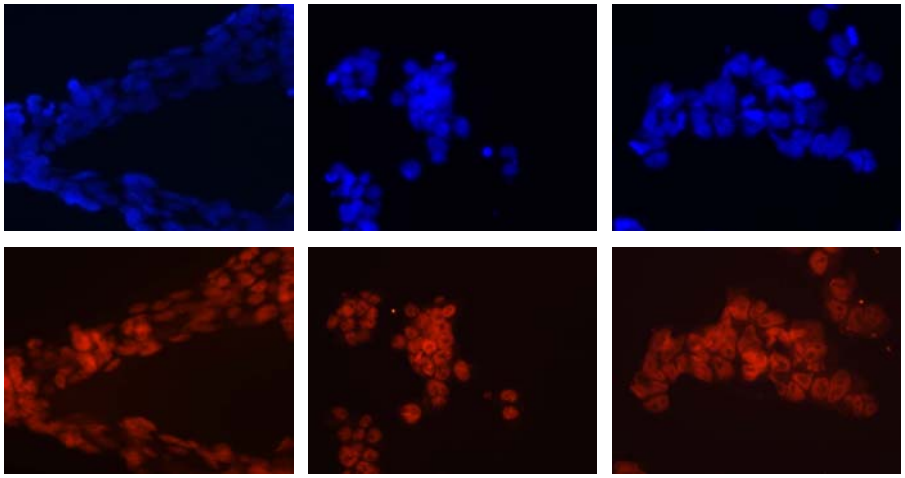


Figure 3.18: Staining of MYC-His-tagged COAS2 protein in the Flp-In system with a Rhodamine-tagged secondary antibody. Top row: Rhodamine/DAPI-stained 293 cells pictured with a DAPI filter, bottom row: identical cells pictured with a Rhodamine filter. Left: constitutive expression, center: parental Flp-In 293 negative control, right: tetracycline induced expression of COAS2-MYC-His after 18 hours.

3.9.2 Filamin-staining of 293 cells

We wanted to visualise the filamin proteins in cells with over-expression of COAS2 to examine whether this ectopic expression could cause a different pattern of filamin organisation or filamin expression. We stained for actin, using phalloidin, an actin-binding reagent, and filamin, using a goat anti-chicken filamin antibody (whole serum). No evident alteration was seen, as shown in figures 3.19 and 3.20.

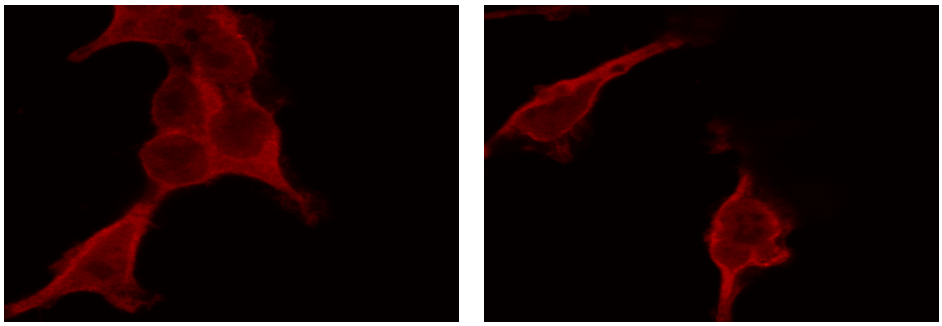


Figure 3.19: *COAS2*-transfected Flp-In cells stained with goat anti chicken filamin antibody (cross-reactive towards human filamins). Left: over-expressed COAS2, right: negative control.

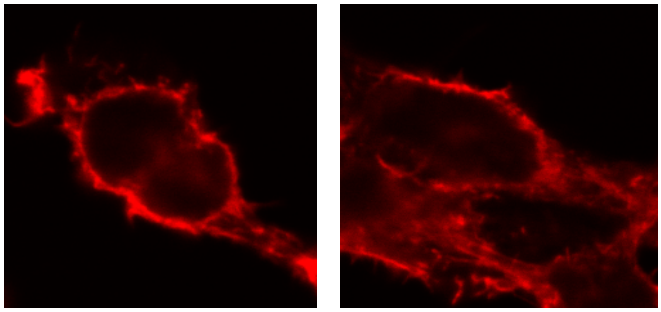


Figure 3.20: *COAS2*-transfected Flp-In cells stained with phalloidin, an actin-binding reagent. Left: over-expressed COAS2, right: negative control.

A specific antibody would facilitate studies of the COAS2 protein. The localisation experiments will thus be repeated upon generation of an antibody towards COAS2.

3.10 Protein expression in bacteria

To produce a monoclonal antibody, a relatively large amount of purified protein is necessary. These amounts are most easily acquired using bacterial expression systems, which give good yield, but lack post-translational processing cellular mechanisms found in mammalian cells. Therefore, expression of a mammalian protein in bacterial cells is not a straightforward procedure, since proteins formed will not be fully processed and thus may cause the formation of inclusion bodies. We chose the BL21(DE3)pLysS cells partly because of their genotype and the tight transcription control mediated by the pLysS function. In this system, the rate of transcription could thus possibly be controlled so as to delimit the formation of inclusion bodies. Two different pET plasmids were used with these cells, to give two differently tagged COAS2 proteins upon expression. The pET30EkLIC plasmid would give a COAS2 protein with S-tag and His_6-tag, whilst the pET3a plasmid only produced a T7-tagged COAS2. All three tags could be used for purification of the recombinant protein, using immunoprecipitation or immunoblotting with appropriate commercibly available antibodies.

The expression of COAS2 in BL21(DE3)pLysS cells was visualized with SDS PAGE and Coomassie staining as described in the methods section. The first gel indicated a product in the pellet fraction of the sample at all time points.

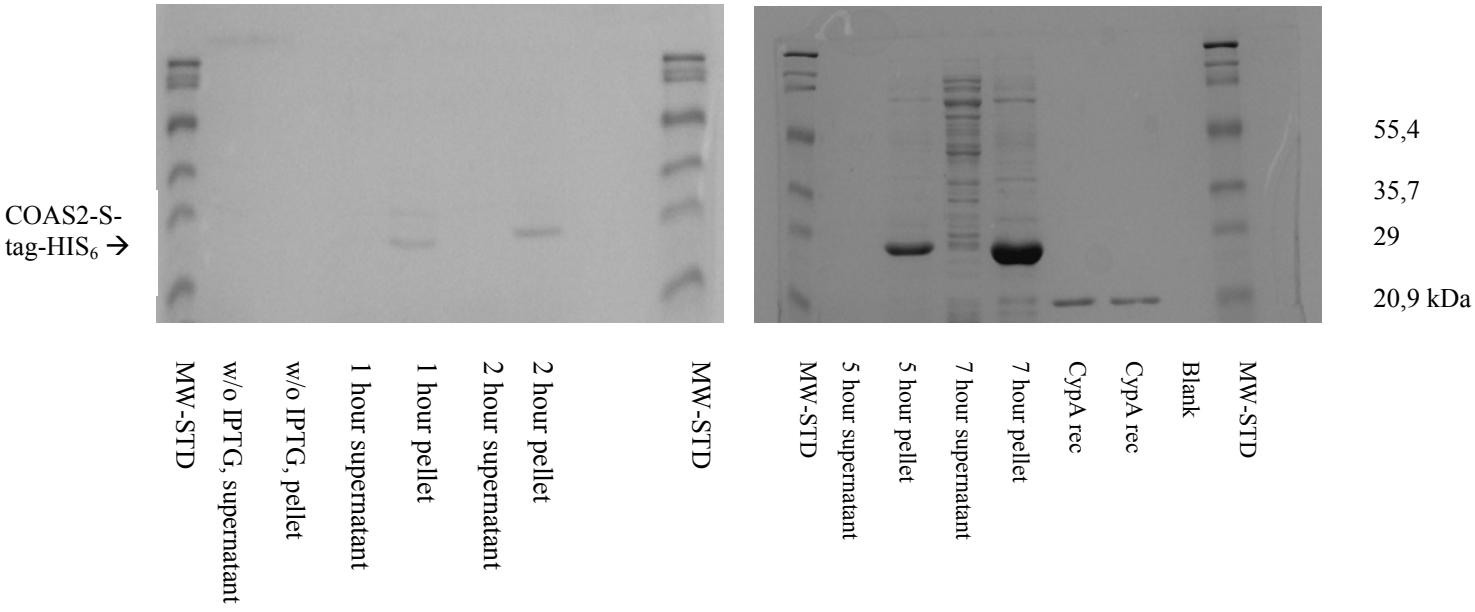


Figure 3.21: SDS-PAGE of time points during COAS2/S-tag-HIS_6 expression in BL21pLysS cells. The 23 kDa protein is found solely in the pellet fraction (migrating as a somewhat larger protein, see figure).

No soluble COAS2/S-tag-HIS$_6$ was formed. Later experiments therefore included additives in the LB in addition to a lowered growth temperature (30 °C) to overcome this problem. A construct producing cyclophilin A, however, produced the expected soluble protein in parallel cultures. We tried LB with added salt (0,02 mM $CaCl_2$, 4 mM NaCl), glucose-rich, (0,06-0,015 M), and glucose-free LB, and lysis with an acidic lysis buffer (NaAc, pH 5.2), to solubilize our protein. Seeing no formation of soluble protein, we repeated the assay at room temperature with the same conditions. After two experiments, the pellet suspension was rinsed with a blend of EGTA and Tween to extract soluble COAS2 as described in section 2.15.1, but the attempt was not successful, as judged from a Coomassie-stained SDS-polyacrylamide gel: (figure 3.22)

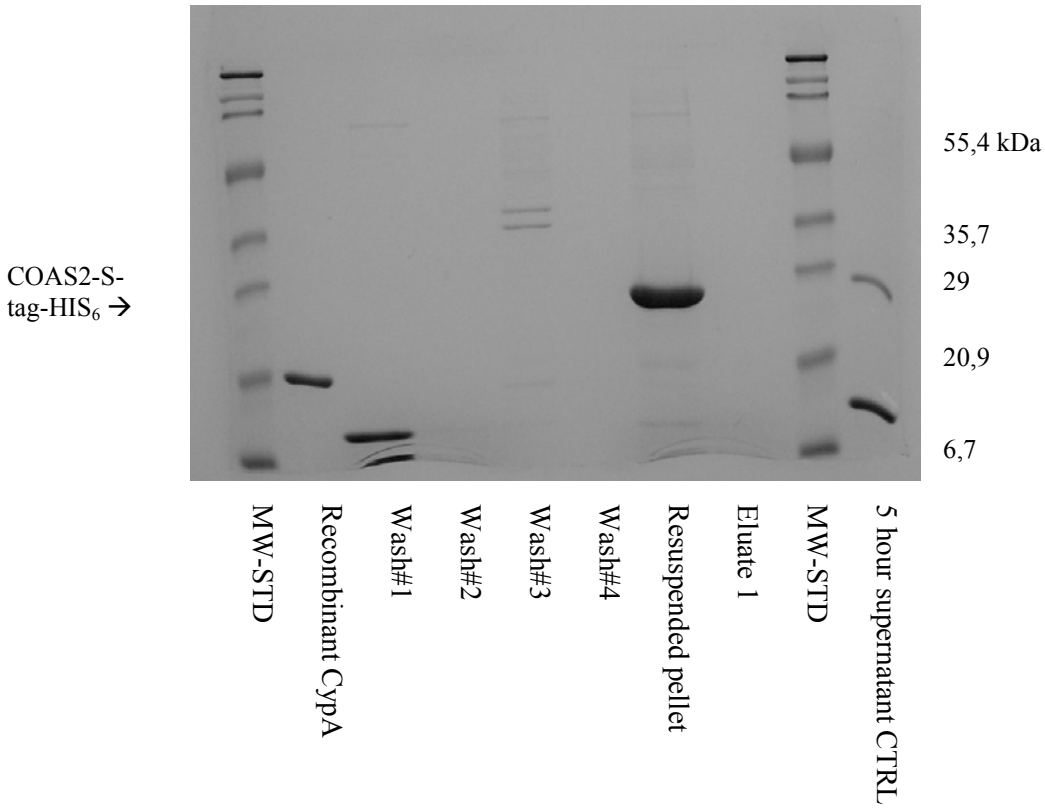


Figure 3.22: SDS-PAGE of time points during COAS2/S-tag-HIS$_6$-expression at 26°C in BL21pLysS cells. The protein remains in the pellet fraction after Tween-assisted purification.

The pET3a/COAS2 construct was applied for production of a T7-tagged COAS2, but even here, only insoluble protein was formed despite altered growth conditions and subsequent purification attempts (results not shown). Our collaborators in the group of Dr. Schiene-Fischer at the Max Planck Institute in Halle, Germany, came to the same conclusions. COAS2

forms insoluble protein in these bacterial systems, and may not be refolded or extracted using detergent washes from the aggregates formed in the pellet fraction.

3.10.1 Protein expression in the RTS100 system

The RTS100 cell free protein translation kit from Roche was used for attempted *in vitro* expression of COAS2. If positive results were achieved, the larger RTS500 system would be the next step, increasing the amount of produced soluble protein. In the RTS100 cell free system, the existence of the egfp control protein was easily verified in UV light, but with more difficulty in a SDS-PA gel. Attempts were made to up-concentrate the His_6-tagged egfp protein using Ni^{2+}-agarose columns were not successful. Several proteins were sought expressed using the RTS100 system, including COAS2-T7, COAS2-S-tag-His_6 and S100A4-His_6 (courtesy of Dr. G. Mælandsmo). In all cases, enough protein was not produced for Coomassie Brilliant-Blue staining. Since the protein lysates created "smears" where COAS2 and other products were difficultly located, we finally attempted to solubilise present protein using a detergent-based procedure described by the supplier: (figure 3.23)

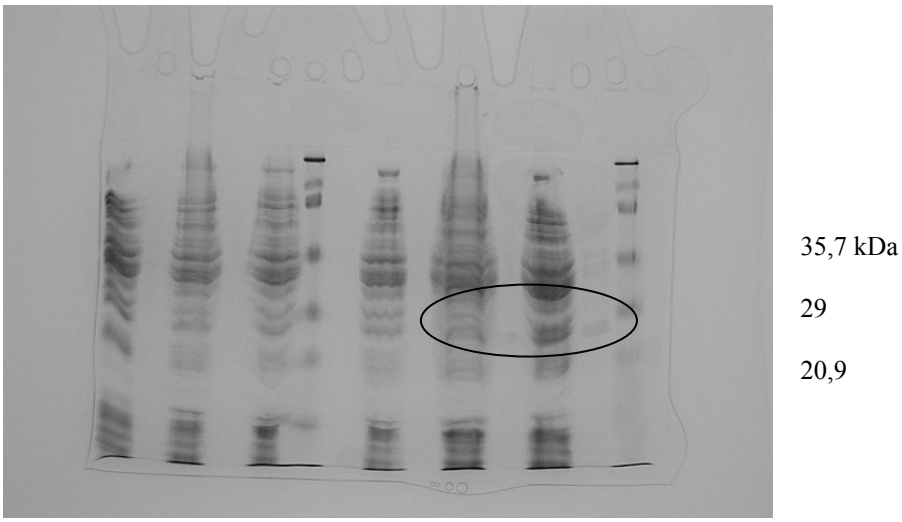


Figure 3.23: SDS-PAGE showing expression levels of COAS2 in the RTS100 cell free system after EGTA-Tween^{20}-purification. The weak egfp positive control proteins (28 kDa) are indicated with an ellipsoid. Lanes number 1,3,5,8,10, 12 and 14 contain non-purified protein lysates. Even numbers contain MW-standards (lanes 6 and 14), purified COAS2-T7 (lanes 2 and 4), purified COAS2-S-tag-HIS_6 (lanes 7 and 9), and egfp controls (lanes 10-13).

Sufficient amounts of protein (for Coomassie Brilliant Blue-staining) were not produced. The pET vectors from Novagen were applied in this assay. We later learned that their backbone was not ideally suited for the RTS system, in contrast to the pIVEX plasmids from Roche. However, we were subsequently offered to express protein for monoclonal antibody production in eukaryotic cells, and thus decided to change expression system from bacteria to insect cells.

3.10.2 Protein expression in insect cells with the Bac-to-Bac system

As the Bac-to-Bac® system was readily established and managed by Dr. David J. Warren, we opted for COAS2-expression in this system, which is based upon baculovirus transduction of insect cells with concomitant protein production. The necessary primary transfer vector was generously offered to us, and the cloning procedure followed the flow chart displayed earlier in this section. Dr. Warren gave us seven different insect cell lysates after harvest and separation of pellet and supernatant. A band of correct size was detected as indicated in figure 3.22, but COAS2 was seemingly present at low levels in the supernatant fraction. Usually, the pellet contains cell wall material and aggregated insoluble protein. Based upon former knowledge with bacterial expression of COAS2, our protein might be present (in high amounts) in this fraction. We did, however, only obtain the supernatant for visualization and quantification of the gene product, and do therefore not know whether more COAS2 was present as insoluble protein in the pellet fraction (potentially as Russell bodies or aggresomes in the insect cells).

Dot blots were performed as outlined in section 2.16.6. Based upon dot blot results, we used 0,5 μg/ml final antibody concentrations for our western assays (the standard varies from 0,1-1,0 μg/ml for most antibodies). In parallel, lysates containing a small HIS_6-tagged protein named CA125 were used as control. Both inserts shared the pFastBacHTa vector backbone. We quantified the total protein content with Bradford reagents, and loaded a range of 2 to 20 μl with 6 x loading buffer onto a 5% + 12% PA gel. The gel was run, blotted and developed according to the methods section, and finally visualized using chemoluminescence.

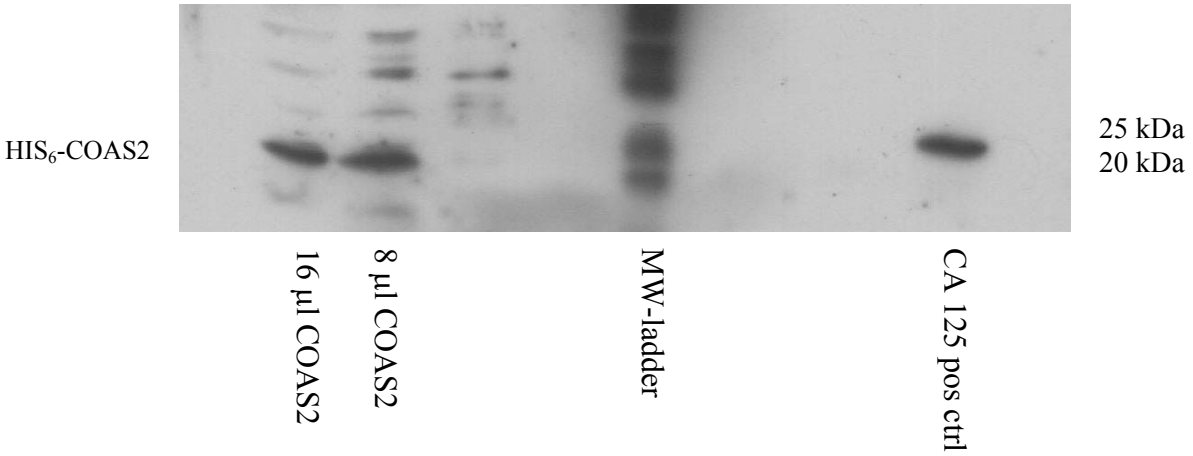


Figure 3.24: Western blot of insect cell lysates, indicating the 21 kDa recombinant COAS2 and the 24 kDa CA125 control protein. The α-His_5 antibody (Qiagen) also bound MW-standard proteins in addition to several unknown probably His-containing proteins in the total cell lysate.

The HIS_6-COAS2 protein showed low, yet even levels of expression in all samples. The CA125 control had higher levels of expression: only 2-3 µl was needed in comparison to 10-12 µl of COAS2-lysate in the above picture. This ratio was later confirmed in parallel blots (not shown). However, the pellet fraction was not examined for existence of the HIS_6-tagged COAS2 protein, where some protein might be present. The lysates have been shipped to our collaborating group at the Max Planck Institute in Halle, Germany, where purification experiments will be performed.

4 Discussion

Throughout this work, attempts have been made to characterise the novel cyclophilin coded by Chromosome One Amplified Sequence 2 (*COAS2*). Amplified and over-expressed in a subset of human tumours of mesenchymal origin, *COAS2* holds great interest as a putative oncogene (Meza-Zepeda et al., 2002). An oncogene can potentially promote cellular growth through deregulation of signal pathways, possibly affecting the cell cycle regulation (Hartwell et al., 1994; Enzinger, 1995; Hesketh, 1997; Ford and Pardee, 1999). In addition, it could, as seen for the homologue PPIA in collaboration with other proteins, confer resistance to chemotherapeutic agents (Hesketh, 1997; Hirai et al., 1999; Kawazu et al., 1999).

As mentioned earlier, COAS2 is a member of cyclophilin family of peptidyl-prolyl cis-trans isomerases (PPIases) (Meza-Zepeda et al., 2002). Their enzymatic activity is thought to be a rate-limiting step during protein folding and may modulate the activity of other proteins and their susceptibility to e.g. phosporylation by bending their peptidyl-prolyl bonds, thus changing their structure (Fischer et al., 1984; Etzkorn et al., 1994; Shaw, 2002). Because all amino acids necessary for this activity are preserved in COAS2, it is thus possible that COAS2 can exert such PPI-activity.

However, because cyclophilins are abundantly expressed and exert diverse functions (Fischer and Schmid, 1990; Schiene-Fischer and Yu, 2001; Bukrinsky, 2002; Shaw, 2002), many possible functional roles could be foreseen for COAS2. To investigate some possible oncogenic function(s), cell growth and cell cycle regulation experiments were performed using stably transfected cell lines with over-expression of COAS2. In parallel, production of the COAS2 protein in model systems for monoclonal antibody production was attempted. Furthermore, we aimed to generate a transcription profile of 293 cells over-expressing COAS2 using cDNA microarray analysis, as this could possibly link the gene to known regulatory cellular pathways. Characterisation of the mammalian cell clones, confirming the proper insertion and expression of the gene product were performed using molecular biology techniques. Later, the project evolved to include confocal microscopy-based localisation studies using cell lines stably transfected with epitope-tagged COAS2.
Of course, the adapted model systems are not ideal for the characterisation of a gene found amplified and over-expressed in sarcomas and breast cancer. They were *a priori* selected due

to ease of manipulation during the time schedule set for this work (traditionally 12-18 months). As mentioned earlier, members of Ola Myklebost's group are currently attempting to establish mesenchymal stem cell line models. These will be of major importance to future research on human sarcomas, including the cellular function(s) of COAS2.

4.1 In silico analysis of the *COAS2* sequence

DNA sequence analysis may often indicate the structure and possible function of the gene product. Hence, domain searches at several online databases have been applied in this work to acquire information regarding *COAS2*. As mentioned, *COAS2* has great amino acid homology to several cyclophilins, with a maximum of 84% amino acid identity to PPIA (Meza-Zepeda et al., 2002). In earlier work done on *COAS2* sequence, Ms. Lygren and Dr. Meza-Zepeda generated hydropathy plots and an mRNA expression profile of the cyclophilin in a panel of tumors and some normal tissues. Later, Dr. S. Tanaka reported it to be upregulated in liver cancer (direct submission to the NCBI database), and termed the gene 'cyclophilin-LC'. Additional mRNA or protein data has until the time of writing not not been published. The early findings indicate that the gene product may exert a similar function as CypA, having conserved the amino acids required for PPI acitivity (Meza-Zepeda et al., 2002). Upon examination of the translated mRNA, the high occurrence of acidic amino acids in COAS2 explains the computed pI of 9.3 compared to the 7.7 of PPIA, but this has not yet been associated with any functional differences. Furthermore, the glycosaminoglycan (GAG) binding sites at amino acids #148-155 of both proteins differ in one amino acid only, an aspect which may indicate that the COAS2 gene product also binds GAG-containing proteins, but perhaps slightly different polycarbohydrates than those of PPIA, whose GAG-site has been shown to bind to heparins (Saphire et al., 1999). COAS2 and PPIA may therefore be expressed during different conditions or at different locations, targeting interaction partners via non-identical hydrophobic or hydrophilic parameters. *COAS2* does not; however, seem to possess other known interaction domains, neither N- nor C-terminally.

Cyclophilin A has many pseudogenes (www.pseudogene.org), and it cannot be excluded that the COAS2 gene could be a pseudogene of CypA. Pseudogenes may be generated by retrotransposition (Haendler and Hofer, 1990; Zhang et al., 2004), i.e. reverse transcription of

a mRNA transcript with subsequent re-integration of the cDNA into the genome. Since *COAS2* does not have any introns, as opposed to CypA, it may have been generated by retrotransposition. Such copies of genes are termed ***processed*** pseudogenes, and when generated by this mechanism, such "retrocopies" of genes may, due to redundancy, accumulate random mutations over the course of evolution, and either acquire new functions, or become processed pseudogenes (Haendler and Hofer, 1990; Zhang et al., 2004). Of course, a requirement for a gene to be termed as a pseudogene is that no protein is produced (Zhang et al., 2004). However, the selection of over-expressed, amplified *COAS2* in tumours, the absence of nonsense mutations and the conservation of the active site argue against this being a pseudogene. In a retrotransposed pseudogene, the promoter sequence may be unefficient or absent (Haendler and Hofer, 1990; Zhang et al., 2004). Again, the fact that *COAS2* is over-expressed and amplified in tumours indicates that an active promoter does exist; alternatively that *COAS2* could be over-expressed and amplified in a subset of tumours as a result of aberrant promoter regulation.

Database annotations suggest that *COAS2* may exist as a longer transcript, where the 5' end is prolonged in respect to the 495 bp gene used in this work. However, the annotations are purely based on sequence analysis and theoretical predictions, and may be incorrect, possibly related to the existence of genomic upstream start codons (for more on database annotation, please see http://www.ncbi.nlm.nih.gov/Web/Newsltr/FallWinter03/builds.html).

Such upstream codons create minicistronic sequences, a common trait thought to have regulatory functions in the human genome (Kozak, 1987). Regardless of computer-based errors, examination of the NM_178230 660 bp sequence reveals a 5'leader sequence of 75 bp harbouring one upstream ATG with a cytosin at –3. This indicates that this upstream coding sequence may function as a regulatory minicistron for the *COAS2* gene (Kozak, 1987). Furthermore, if expressed, this small coding sequence should theoretically pose no problem for *in vitro* synthesis (e.g. leaky scanning, resulting in ribosome throw-off) (Kozak, 1987). Interestingly, the Kozak-sequences upstream of the start codons in *COAS2* and *PPIA* are 5' CTATCAGCC 3' and 5' CTAGTAGCC 3', respectively. These are quite different when compared to the consensus sequence (GCC) GCCRCCATGG (where R = A or G), which often is recommended in system manuals for molecular cloning procedures. But they are relatively similar to each other, indicating that COAS2 mRNA could be translated, however possibly at lower levels than that of PPIA due to the the lack of stabilising guanidines (Kozak,

1987) at positions -6 and -9 in the COAS2 Kozak sequence. Indeed, the highly expressed PPIA has a stabilising G at –6.

In conclusion, examinations of DNA, mRNA and amino acid sequences indicate that the COAS2 protein has a pI of 9.3, and a C-terminal proteoglycan-binding sequence with 87,5 % conserved amino acids compared to that of PPIA. In addition, earlier evidence suggests that all the amino acids necessary for PPI enzymatic activity in PPIA are preserved in COAS2. Moreover, divergence of the two sequences through evolution or retrotransposition may have created a more polar and hydrophilic cyclophilin CypA-homologue such as COAS2, with complementary functions in- or outside the eukaryote cell. Furthermore, COAS2 is amplified and over-expressed in a subset of human tumours of mesenchymal origin, indicating that it is not a processed pseudogene. Finally, the existence of the COAS2 protein has not yet been reported, but comparisons of PPIA and COAS2 show that the –10 to +1 upstream sequence of PPIA has been preserved, suggesting that the COAS2 mRNA is translated in human cells. Based on the sequence analysis, it may be advisable to include the 5' sequences reported in new constructs, and perhaps also to change the Kozak sequence to that of PPIA, which works well in our hands.

4.2 Creation and characterisation of stably transfected Flp-In cell lines

COAS2 had earlier been studied using transient expression only (Meza-Zepeda et al., 2002). A cell line with a stably integrated single copy of a gene of interest (GOI) downstream of a controllable Tet-promoter as described by (O'Gorman et al., 1991) and (Yao et al., 1998), would be a better tool than transient transfection methods for functional gene characterisation. First, this system would imply better control of expressed protein, in contrast to the uncontrolled expression levels directed by strong promoters when transient transfection is applied. Second, intracellular aggregation of proteins and possible downstream artefacts, as experienced in the previous studies, could thus putatively be avoided (Sauer, 1994; Yao et al., 1998). A system where stable transfectants are created would thus reduce the number of potential problems (Sauer, 1994). In this work, the Flp-In system was applied to create such cell lines; a system based on antibiotic selection of positive clones harbouring the GOI, integrated at a specific site (Craig, 1988; Sauer, 1994).

Although COAS2 is a member of a protein family with numerous and diverse functions, its possible effects on cancer-related mechanisms of growth and cell cycle regulation were prioritised in this work due to the proposed oncogenic role. Microarray analysis (of cells over-expressing COAS2) could however give clues to downstream effects and link the protein to any type of signal pathway.

Upon creation of the Flp-In cell lines used in this work, the reading frame of the integration site was altered towards expression of a Hygromycin-resistance molecule after integration of the GOI. This should ideally fully halt the transcription of the fusion gene *lacZ*Zeo, inevitably making the cells sensitive to Zeocin, and resistant to Hygromycin (Craig, 1988; Sauer, 1994). Very few clones did exhibit Zeocin-sensitivity when tested, a fact that still has not been explained. Interestingly, the levels of *COAS2* were shown to differ between Zeocin-resistant and Zeocin-sensitive clones, as indicated in figure 3.7. To rule out that this could be due to different amounts of integrated COAS2, we determined the copy numbers using a Southern blot, as seen in figure 3.5. When insertion occurs at the single FRT site, it should do so only once, but an option may be either duplicate insertion, i.e. the insertion of the GOI-containing plasmid twice or more in an existing FRT site, insertion of one or more copies at other location(s), or (in theory) multiple copies of FRT sites. However, when examined, both parental and *COAS2*-transfected cell lines seemed to have one copy of the *lacZ*-Zeo fusion gene in their genome, and only the transfected cells had one copy of the Hygromycin gene, as expected after integration of only one copy of *COAS2*.

The supplier of the Flp-In system offers a third test for selection of positive clones, but this has not been used in this work. This test benefits from the parental cell line-expression of the *E. coli lacZ*Zeo gene, the *lacZ* moiety coding for the enzyme β-galactosidase. This enzyme of bacterial origin can utilize the synthetic substrate X-gal to yield blue-stained cells (Mensa-Wilmot and Englund, 1992). In positive clones, the transcription of the lacZ gene is disrupted, and the cells will therefore not stain blue when fed X-gal to the cell growth medium (blue-white selection method known from *E. coli*). To apply this method for initial clonal screening instead of screening for Zeocin-sensitivity may be advisable for future purposes, since it is less time consuming. This could improve the selection of positive clones.

Interestingly, expression levels of *COAS2* in the transfected Flp-In cell lines have been shown to vary between Zeocin-sensitive and Zeocin-resistant clones. However, promoter release by

tetracycline addition to the cell growth medium generates an induction of *COAS2*-expression in all cases, as seen in figures 3.7 and 3.9, and it thus does not seem that Zeocin screening is a reliable assay for integration.

For further characterization of the FRT site in the cell line genome before and after transfection, one could perform a repeated Southern experiment where a second restriction enzyme is used, in addition to quantification of mRNA (with RT-PCR or titration curves of *COAS2* mRNA levels using northern blots), and protein (with western blotting upon generation of an antibody towards COAS2). At the moment, further information on eventual interclonal differences is thus limited.

4.3 Cell growth assays

Oncogenic activity may be related to mediation of cell proliferation, (Hesketh, 1997), and effects of cyclophilins on cell growth have been reported (Andersen et al., 2003). In addition, the Pin1 protein, a member of the immunophilin family of proteins, has been shown to be a cell cycle regulator (Wulf et al., 2002).

Accordingly, the putative effects of over-expressed COAS2 on cell growth were therefore addressed in this work through the use of several *in vitro* and *in vivo* growth assays.

It is also known that the class A cyclophilins bind the immunosuppressive drug cyclosporin A (CsA) (Fischer et al., 1989). Interestingly, the multidrug resistance pump P-glycoprotein (PGP) has been shown to be noncompetitively inhibited by CsA (Wigler, 1999). In repeated experiments with PGP and cyclosporins, the influence of cellular cyclophilin has been indicated (Fakata et al., 1998; Kawazu et al., 1999; Tiberghien et al., 2000). This suggests a possible role for COAS2 as mediatior of drug resistance, however depending upon the nature of the applied cytostatic. Thus, the ability of COAS2 to confer resistance to three types of cytostatica in this model system was examined.

In vitro growth rate, determined as a part of the cytostatics testing, showed that cells over-expressing COAS2 grow slightly faster than their parental cell lines (figure 3.11). Moreover, the two soft agar assays indicated a moderate but reproducible increased growth potential for the *COAS2*-transfected cells with constitutive expression when compared to the parental Flp-

In 293 cell line. For *in vivo* confirmation, cells were subsequently injected subcutaneously into nude mice, and the following xenograft growth experiments performed by Mr. Christoph Müller supported the soft agar data (figure 3.13). Furthermore, as expected from a putative oncogene, over-expression of COAS2 in these stable cell lines does not induce apoptosis. Over-expression of *COAS2* did not; however, seem to exert a large effect on cell growth of 293 Flp-In cells, although the results in all systems were coherent, suggesting a positive effect. Further analysis is nevertheless needed to confirm that these results are not just due to clonal variation. It should be noted that even a moderate growth advantage might be selected for in a tumour over years of development, and could even be increased under certain *in vivo* conditions.

We used flow cytometry to investigate effects on cell cycle distribution in cells with over-expressed COAS2. However, such effects are often masked when using over-expression of proteins in *in vitro* assays. They may only be seen under conditions of low cell growth, or only towards a certain background of defects (e.g. in "knockout" or "knockdown" cells). Therefore, experiments should ideally be performed under conditions of reduced growth (e.g. low fetal calf serum (FCS) concentration in cell growth medium), or with colchicine-synchronised cells. We thus compared the cell cycle distribution of Flp-In 293 cells with and without over-expression of COAS2, with different levels of (FCS) in the cell growth medium. From our series of experiments, the COAS2 protein apparently does not seem to have a gross effect on cell cycle regulation, neither when cells were grown in 2 % FCS nor in 8 % FCS. Interestingly, a small peak occurred in control samples without over-expression of COAS2 (figure 3.15 D). It is possible that these cells are apoptotic cells, and that over-expressed COAS2 counteracts this apoptosis, since this was not observed in samples with over-expressed COAS2. Furthermore, this would have significance for cell growth in cells with over-expressed COAS2; the lack of apoptosis could explain the moderately increased growth seen in the model cell lines used in this work. Moreover, a continued negative regulation of apoptosis will affect tumourigenesis, strengthening the role of COAS2 as amplified and over-expressed in a subset of human cancers. This possible inhibition of apoptosis was however insufficient to counteract the cytostatic drugs analysed in other experiments (see further on).
Consequently, future experiments where the serum levels are kept below 1 % in addition to synchronisation experiments and measurement of apoptosis may give more conclusive results to a possible role of COAS2 in cell cycle regulation. Preliminary results suggest that COAS2

does not affect cell cycle regulation when overexpressed in the model system used in this work. This is of course when over-expressed alone; its regulatory function may be depending on other interaction partners not present in our cell system.

4.4 Resistance to cytostatics

Investigations of cytostatic response using SRB colorimetric assays showed that, regardless of their type, the applied cytostatica caused both controls and samples to react equally, as shown in figure 3.14. The *COAS2* gene product may thus not confer resistance to cisplatin, methotrexate nor doxorubicin, at least not in this system. However, the importance having the correct model system should be emphasized; important experiments will be repeated when a mesenchymal model cell line has been established.

4.5 Transcription analysis of cells over-expressing COAS2

Being able to monitor expression changes in a large set of samples is the major advantage of microarray analysis (Ball et al., 2002; Miller et al., 2002; Quackenbush, 2002). The microarray technology is now an established method for transcription analysis. In Norway, the activity is guided by the Norwegian Microarray Consortium (NMC), a national collaboration between universities in Oslo, Bergen and Trondheim. 15k spotted cDNA arrays from the NMC have been used in this work to identify possible effects on the transcriptome by our cyclophilin.

Although it is unlikely that COAS2 participates directly in transcription regulation, perturbing the cell by COAS2 over-expression might affect its transcriptional profile, indirectly linking the COAS2 protein to signal pathways. Therefore, we used microarray analysis on stably transfected cell lines where over-expression of COAS2 was achieved from a controllable TetOn-promoter (Craig, 1988). Time points were selected to examine possible primary (early) and secondary (late) effects of over-expression. The high expression of a *COAS2* transcript upon tetracycline induction was illustrated in figures 3.7 and 3.8. The *COAS2* cDNA was not present on the arrays, as opposed to other cyclophilins (*PPIA*, *PPIC* and *PPIF*), to whom COAS2 displays a high degree of amino acid homology. The other cyclophilins were not

found to be highly upregulated (maximum sample/reference ratio of 1.4 for the PPIA, as shown in table 3.2). This suggests that the stringency of our experiments was sufficient, minimising the degree of cross-hybridisation from other cyclophilin cDNAs, further strengthening the obtained results.

Upon examination of the data, two Sry-box (SOX) transcription factors were found to be upregulated in the cell line material, as seen in table 3.2. These two were interesting because the SOX family of proteins are involved in many stages of cell differentiation, as controllers of embryonal neurogenesis, (Muhr, J. et al, in press), chondrogenesis and osteogenesis (Zhao et al., 1997; de Crombrugghe et al., 2000; de Crombrugghe et al., 2001; Lefebvre et al., 2001). This could represent a link between COAS2, which has been found to display high mRNA levels in osteosarcoma samples with amplification (Meza-Zepeda et al, 2002), and chondro- and osteogenesis. In fact, 1q amplification has been connected to ossifications and osteosarcomatous differentiations in earlier work done on well-differentiated liposarcomas (WDLPS) (Forus et al., 2001b).

Together with a known modulator of mammalian osteogenesis and chondrogenesis, *Sox9*, the L-*Sox5* and *Sox6* have been shown to regulate collagen type two alpha one (col2a1),(Zhao et al., 1997; Lefebvre et al., 1998; Smits et al., 2001). However, COL2A1 was present on the cDNA arrays, but only found weakly upregulated in one sample, and the control hybridisation performed using northern analysis supported the microarray data.

On the other hand, among the upregulated genes in our data was the collagen type 1 alpha 2, which normally is present at later stages of bone differentiation (Lefebvre et al., 2001). This was expressed at detectable levels in few samples, and has not been further investigated. However, the connection between COAS2 and cartilage- and bone-expressed proteins is nevertheless very interesting, and will be examined further.

Verification of high expression of L-SOX5, SOX6 and SOX9 was attempted using northern blotting (SOX9 despite the fact that no *SOX9* cDNAs were present on the cDNA array slides). This was, however, not successful, probably because the levels were below the detection limits of the respective probes.

Verification using northern was finally abandoned as oligonucleotide arrays became available, arguing that results from the oligonucleotide arrays presumably would be able to confirm cDNA array data. However, we ran out of RNA, making repeats of the cDNA array experiments with oligonucleotide arrays difficult. Evidently, all interesting genes must be examined for expression and thus confirmed as high- or upregulated using northern (or potentially western blotting), or rather real time reverse transcriptase PCR upon repetition of the experiment. Without experimental confirmation, the link between COAS2 and the above genes thus remains interesting, but not confirmed at the time of writing.

4.6 Cell biological studies

Localisation of a protein may give valuable clues to its function (Jarvik and Telmer, 1998). The cyclophilin family, subclass A, to which COAS2 belongs, are ubiquitously expressed in human tissues (Bukrinsky, 2002; Galat, 2003). PPIA, to which COAS2 shares 84 % amino acid similarity, is preferentially located in the cytoplasma (Shaw, 2002), but has also been shown to be present outside the cell membrane when active during HIV-virion formation (Saphire et al., 1999). In addition, localisation of the endogenous protein demands a highly specific antibody, but this is not yet available. We could only therefore do *in vitro* experiments using epitope-tagged COAS2 in this work.

Thus, three tags were applied to *COAS2* through molecular cloning, either *egfp* or *MYC-HIS*$_6$. All constructs were transfected into the Flp-In system to establish stable integrants under tetracycline-controlled TetOn promoters (O'Gorman et al., 1991).

The egfp protein will fluoresce when excited under blue light, and is therefore an ideal target protein, making fixation and thus destruction of intracellular networks unnecessary (Tsien, 1998). Live imaging using confocal microscopy was planned for the egfp fusion proteins, whilst the MYC-HIS$_6$-tag was intended for use in fixed-cell imaging and possible pulldown experiments.

4.6.1 Localisation of the COAS2-MYC-HIS_6 protein

COAS2 mRNA was identified in all samples using northern blot analysis (figures 3.9 and 3.10), and the sequences were all verified by sequencing. However, when interpreting the levels of *COAS2* over-expression (figure 3.7 and 3.9), it was noted that in samples with no over-expressed *COAS2*, a band was present. This could be due to endogenously expressed *COAS2* in the samples. To more precisely investigate the level of expression of the uninduced transgene, probes containing the *MYC-HIS_6* tag only may be hybridised to the northern filter shown in figure 3.9.

Detection of COAS2-MYC-HIS_6 was thus subsequently attempted using western blotting. However, using antibodies towards both tags, proteins were not detected in the cell lysates under neither denaturating nor native conditions despite the fact that both antibodies could detect respective control MYC and HIS_6-proteins, regardless whether the tag was located N- or C-terminally. This phenomenon remains unexplained, because if the (MYC-HIS_6) tag was properly translated and the (COAS2-MYC-HIS_6) protein incorrectly post-translationally processed or folded, one would still be able to identify the protein in a western blot with properly-functioning antibodies towards the tags (Harlow, 1988). Nevertheless, several possible explanations exist; perhaps the cause could be the lack of laboratory skill or knowledge, i.e. that the protein is present at very low levels, but we were not able to fully optimise our detection system. Alternatively, the mRNA may be incompletely processed post-transcriptionally, regulating or fully inhibiting translation.

Attempts were however further made to localise COAS2-MYC-HIS_6 with confocal fluorescence microscopy. As shown in figures 3.17 and 3.18, the obtained signals were not specific for COAS2, as a nuclear staining excluding the nucleoli was seen for all samples, including the negative control. Due to the low signal intensity of the His_5 antibody, and the absence of a signal when stained with the MYC antibody, we were therefore not able to localise the MYC-HIS_6-tagged COAS2 protein in the cells.

Here too, several explanations may exist. Optimisation of the system was attempted, and extra precautions were taken during the experiments to avoid bleaching of the Rhodamine-conjugated antibody. As mentioned earlier, positive induction of COAS2 mRNA upon tetracycline addition was verified (figure 3.9). Nevertheless, without proven existence of the

full length tagged mRNA, western blot- and localisation results will remain unexplained. We must therefore hybridise the northern filter depicted in figure 3.9 with a MYC-HIS_6-probe. If detected, the (*COAS2-MYC-HIS_6*) mRNA should theoretically be translated into a recombinant protein that might be seen on an optimised western blot (Harlow, 1988). However, results indicate that COAS2-MYC-HIS_6 may not be translated correctly, eventually misfolded or aggregated. It can be difficult to assign this to either aggresome- or Russell body formation, since there are no easy ways of identifying these molecular complexes in the cell (Kopito and Sitia, 2000). With electron microscopy, aggregating structures could possibly be seen intracellularly. This was however not possible within the timeframe of this work.

4.6.2 Localisation of COAS2 using egfp fusion proteins

In transient localisation experiments performed by Ms. Lygren, COAS2 in fusion with egfp showed a speckled pattern of intracellular localization in fibroblasts, neither localised to nucleus, endosomes nor lysosomes (Lygren, 2002). This phenomenon could represent a translational error or an artefact of unknown origin. Moreover, Ms. Lygren reported a low level of cells with (<1%) of speckled fluorescence, whereas the levels of fluorescence were noticeably higher in the 40 % of cells transiently transfected with egfp only. This could possibly be a result of interactions between the COAS2-egfp fusion protein and the cytoplasmic proteasomes, recruiting the fusion protein for degradation (see further on). A second explanation may be COAS2-induced misfolding of egfp. Alternatively, COAS2 may have an isomerase activity that prevents proper egfp folding, thereby inactivating the fusion protein's ability to fluoresce.

A possible connection between the cytoplasmic protein degradation complexes known as proteasomes and COAS2 has been reported (Lygren, 2002). Links between COAS2 and the proteasome subunit alpha seven as well as the filamin A and C proteins and COAS2 were established by Ms. Lygren using the yeast-two hybrid method. The proteasomes can be stained using specific antibodies, but will only give an even fluorescence located to the cytoplasmic region of the cells (Dr. H. Stenmark, personal communication).

To further investigate Ms. Lygren's results, Mr. M. Skårn transiently transfected U2OS cells with *egfp*-tagged *COAS2* (figure 3.16). The transiently transfected cells produced a pattern of egfp fluorescence only weakly reminiscent of the speckled pattern seen in Ms. Lygren's

transfected fibroblasts (results not shown). The levels of fluorescence were low when compared to the pEGFP-N1-transfected control, but stained perinuclear substructures looking more like intracellular membranes than the previously observed speckles. Furthermore, the transfected cells displayed even fluorescence (of both COAS2-egfp and egfp-COAS2) also in the cell membrane, a result not seen in earlier experiments. One possibility could be that this is due to binding via the GAG-domain to the cell surface.

Flp-In cell lines with stably integrated egfp-tagged proteins were created. Interestingly, the transfectants expressing the same *egfp*/*COAS2* or *COAS2*/*egfp constructs* did not fluoresce above background levels. This could indicate that the lower, more physiological levels of expression aimed for in this work, created insufficient protein for detection. (The positive control, a cell line over-expressing the egfp protein only, fluoresced brightly when examinated.) It is known that cellular autofluorescence can mask low levels of fluorescent protein (Lichtman, 1994), and we will not therefore exclude a low expression of COAS2-egfp or egfp-COAS2 in these cells.

In light of these results, the egfp fusion proteins expressed from the Flp-In 293 cells are thus most likely functional, as suggested by the transient experiments, but the levels produced from a single copy integrated transgene may be insufficient for detection by fluorescence microscopy.

4.6.3 Altered filamin structures

In light of the yeast two-hybrid results where an interaction between COAS2 and filamins were reported, we tried to examine changes of cellular filamin- and actin structure upon over-expression of COAS2, using confocal fluorescence microscopy. (Actin filaments are cross-linked in e.g. a muscle cell by the filamin proteins; the proteins are thus interdependent.) However, the filamin structures were no different with or without over-expressed COAS2 in the *COAS2*-transfected clone #3. Neither were the actin structures particularly altered as displayed by the levels of the actin-binding reagent phalloidin (figures 3.19-20). Consequently, although only duplicate experiments were performed, over-expressed COAS2 does not seem to have an effect of neither actin- nor filamin structure nor filamin- or actin organisation in this system.

4.7 Protein expression in bacteria and insect cells

The existence of an endogenous COAS2 protein can be addressed with the use of a highly specific antibody towards COAS2.

Moreover, the results found through the use of cell line material must be confirmed with protein data. Purified COAS2 will therefore be used for monoclonal antibody formation in addition to other downstream experiments. Being a member of the cyclophilin family of proteins, which displays great internal sequence similarity (Willenbrink et al., 1995; Galat, 2003), may cause some problems. However, the high pI of COAS2 compared to that of the other cyclophilins suggests that the more polar amino acid sequence of COAS2 could induce an antibody with a high degree of specificity.

The attempts of protein expression in BL21pLysS cells were halted due to formation of insoluble protein (figures 3.21 and 3.22), as affirmed by our more expert collaborators at the Max Planck Institut für Enzymologie in Halle, Germany. Their refolding experiments produced an incorrectly folded COAS2 protein without enzymatic activity in solution, and they were unable to refold it (results not shown). Considering that neither important co- and post-translational modifications nor disulfide bridge formation occur in bacterial cells (Kopito and Sitia, 2000; Sambrook, 2001), it is very likely that insoluble protein aggregates were formed because of incorrect folding or modification errors upon IPTG-assisted induction of protein expression. Even though the formation of aggressomes and inclusion bodies in eu- and prokaryotic model systems is known, its cause has not at present been fully determined (Kopito and Sitia, 2000). However, to address whether inclusion bodies or aggressomes are formed in the cells upon over-expression, staining procedures or electron microscopy must be applied. This was outside the scope of this work.

In most cases, protein researchers find the change of model system to be a better adaptation than refolding experiments (Dr. D. J. Warren, personal communication). Consequently, the production of COAS2 in a eukaryotic model system would be better for all downstream purposes, also because recombinant protein from *E. coli* would most likely harbour lipopolysaccharide contamination, which again could profoundly affect some cellular responses (Sambrook, 2001; Simpson, 2003). Insect cells harbour most of the necessary transcription and translation machinery needed for immaculate synthesis of mammalian

proteins *in vitro* (Jarvis et al., 1990; Wickham et al., 1992), and was chosen for further attempts at COAS2-production.

The viral-based Bac-to-Bac system used for COAS2-expression caused expectations to rise when western blotting identified a band of proper size in the first cell lysates to be analysed (figure 3.24). However, the levels of protein were low, as described in section 3.10.2. Moreover, it should be mentioned that the fraction containing cell debris and possible protein aggregates was not kept and analysed. Attempts were made in order to up-concentrate the His-tagged protein, using immunoprecipitation with Protein G agarose (results not shown). The experiments were repeated several times, under different conditions taken from three different protocols, but the attempts were not successful. Some lysate was therefore subsequently shipped to Germany for analysis, where members of Dr Schiene-Fischer's group will attempt to purify the recombinant protein.

4.8 Future aims

To raise a monoclonal antibody towards COAS2*,* the *COAS2* ORF will either be subcloned into the pET43.1 Nus-tag plasmid for expression in BL21pLysS bacterial cells, or preferentially prepared for protein expression in a yeast (*P. pastoris)* system (in collaboration with Dr. David J. Warren).

Further validation is needed for our *COAS2-MYC-HIS$_6$* construct. Western and northern blotting will be applied to address the expression of COAS2-MYC-HIS$_6$ and COAS2/egfp fusion proteins at mRNA and protein levels. In addition, the existing transfected Flp-In 293 cells expressing COAS2- and epitope-tagged COAS2 may be studied further with confocal microscopy and markers for various cellular substructures for all experiments, or with electron microscopy to see if the proteins are membrane-associated, or form aggregates upon over-expression.

For all experiments, a mesenchymal cell type will be used, most likely from stroma-derived MSC tert 20 (Simonsen et al., 2002), which is currently used in our lab.

Transcriptional profiling (microarray) experiments should be repeated with oligonucleotide arrays, with 4-6 time points for increased clarity. Positive transcripts need, of course, to be

validated, e.g. with quantitative reverse transcriptase PCR. Further characterisation of the promoter sequence of COAS2 will shed light on its pattern of translation *in vivo*, e.g. by using luciferase constructs of *COAS2*.

Further analyses of the interaction partners detected by yeast two-hybrid analysis will be important. This might be accomplished with immunoprecipitation assays using both candidate interaction partners, or by validating the interactions in relevant cells using e.g. peroxisome-directing tags (Nilsen et al., 2004). In addition, the effect of a proteasome inhibitor "cocktail" on expression rates of COAS2 in the Flp-In clones will hopefully be elucidating.

Finally, possible effects of over-expressed COAS2 on cell cycle regulation will be examined using cells grown in conditions of low (<1%) fetal calf serum, following a procedure that has been described earlier in this work. The possible effect(s) of COAS2 on cell death in this system will be further investigated.

References

Abramson, D. H., Ellsworth, R. M., Kitchin, F. D. and Tung, G. (1984). Second nonocular tumors in retinoblastoma survivors. Are they radiation-induced? Ophthalmology **91**(11): 1351-5.

Alberts, B. e. a. (1994). Molecular Biology of the Cell, Garland publishing.

Alitalo, K. and Schwab, M. (1986). Oncogene amplification in tumor cells. Adv Cancer Res **47**: 235-81.

Alitalo, K., Winqvist, R., Lin, C. C., de la Chapelle, A., Schwab, M. and Bishop, J. M. (1984). Aberrant expression of an amplified c-myb oncogene in two cell lines from a colon carcinoma. Proc Natl Acad Sci U S A **81**(14): 4534-8.

Allain, F., Vanpouille, C., Carpentier, M., Slomianny, M. C., Durieux, S. and Spik, G. (2002). Interaction with glycosaminoglycans is required for cyclophilin B to trigger integrin-mediated adhesion of peripheral blood T lymphocytes to extracellular matrix. Proc Natl Acad Sci U S A **99**(5): 2714-9.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. J Mol Biol **215**(3): 403-10.

Andersen, H., Jensen, O. N. and Eriksen, E. F. (2003). A proteome study of secreted prostatic factors affecting osteoblastic activity: identification and characterisation of cyclophilin A. Eur J Cancer **39**(7): 989-95.

Anderson, M., Fair, K., Amero, S., Nelson, S., Harte, P. J. and Diaz, M. O. (2002). A new family of cyclophilins with an RNA recognition motif that interact with members of the trx/MLL protein family in Drosophila and human cells. Dev Genes Evol **212**(3): 107-13.

Aramburu, J., Rao, A. and Klee, C. B. (2000). Calcineurin: from structure to function. Curr Top Cell Regul **36**: 237-95.

Arevalo-Rodriguez, M., Cardenas, M. E., Wu, X., Hanes, S. D. and Heitman, J. (2000). Cyclophilin A and Ess1 interact with and regulate silencing by the Sin3-Rpd3 histone deacetylase. Embo J **19**(14): 3739-49.

Aviv, H. and Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc Natl Acad Sci U S A **69**(6): 1408-12.

Avramut, M. and Achim, C. L. (2003). Immunophilins in nervous system degeneration and regeneration. Curr Top Med Chem **3**(12): 1376-82.

Ball, C. A., Sherlock, G., Parkinson, H., Rocca-Sera, P., Brooksbank, C., Causton, H. C., Cavalieri, D., Gaasterland, T., Hingamp, P., Holstege, F., Ringwald, M., Spellman, P., Stoeckert, C. J., Jr., Stewart, J. E., Taylor, R., Brazma, A. and Quackenbush, J. (2002). Standards for microarray data. Science **298**(5593): 539.

Bartek, J. and Lukas, J. (2001). Are all cancer genes equal? Nature **411**(6841): 1001-2.

Basu, A., Das, M., Qanungo, S., Fan, X. J., DuBois, G. and Haldar, S. (2002). Proteasomal degradation of human peptidyl prolyl isomerase pin1-pointing phospho Bcl2 toward dephosphorylation. Neoplasia **4**(3): 218-27.

Baum, C., Forster, P., Hegewisch-Becker, S. and Harbers, K. (1994). An optimized electroporation protocol applicable to a wide range of cell lines. Biotechniques **17**(6): 1058-62.

Berwick, M. and Vineis, P. (2000). Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. J Natl Cancer Inst **92**(11): 874-97.

Biedler, J. L. and Spengler, B. A. (1976). Metaphase chromosome anomaly: association with drug resistance and cell-specific products. Science **191**(4223): 185-7.

Borel, J. F., Feurer, C., Gubler, H. U. and Stahelin, H. (1976). Biological effects of cyclosporin A: a new antilymphocytic agent. Agents Actions **6**(4): 468-75.

Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D. and Kuriyan, J. (1998). The structural basis of the activation of Ras by Sos. Nature **394**(6691): 337-43.

Brazin, K. N., Mallis, R. J., Fulton, D. B. and Andreotti, A. H. (2002). Regulation of the tyrosine kinase Itk by the peptidyl-prolyl isomerase cyclophilin A. Proc Natl Acad Sci U S A **99**(4): 1899-904.

Brizel, D. M., Scully, S. P., Harrelson, J. M., Layfield, L. J., Bean, J. M., Prosnitz, L. R. and Dewhirst, M. W. (1996). Tumor oxygenation predicts for the likelihood of distant metastases in human soft tissue sarcoma. Cancer Res **56**(5): 941-3.

Broach, J. R., Guarascio, V. R. and Jayaram, M. (1982). Recombination within the yeast plasmid 2mu circle is site-specific. Cell **29**(1): 227-34.

Bruns, K., Fossen, T., Wray, V., Henklein, P., Tessmer, U. and Schubert, U. (2003). Structural Characterization of the HIV-1 Vpr N Terminus: EVIDENCE OF cis/trans-PROLINE ISOMERISM. J. Biol. Chem. **278**(44): 43188-43201.

Bukrinsky, M. I. (2002). Cyclophilins: unexpected messengers in intercellular communications. Trends Immunol **23**(7): 323-5.

Carpentier, M., Allain, F., Slomianny, M. C., Durieux, S., Vanpouille, C., Haendler, B. and Spik, G. (2002). Receptor type I and type II binding regions and the peptidyl-prolyl isomerase site of cyclophilin B are required for enhancement of T-lymphocyte adhesion to fibronectin. Biochemistry **41**(16): 5222-9.
Cavenee, W. K. and White, R. L. (1995). The genetic basis of cancer. Sci Am **272**(3): 72-9.
Courtenay, V. D. and Mills, J. (1978). An in vitro colony assay for human tumours grown in immune-suppressed mice and treated in vivo with cytotoxic agents. Br J Cancer **37**(2): 261-8.
Craig, N. L. (1988). The mechanism of conservative site-specific recombination. Annu Rev Genet **22**: 77-105.
Cui, Y., Mirkia, K., Florence Fu, Y. H., Zhu, L., Yokoyama, K. K. and Chiu, R. (2002). Interaction of the retinoblastoma gene product, RB, with cyclophilin A negatively affects cyclosporin-inhibited NFAT signaling. J Cell Biochem **86**(4): 630-41.
Dal Cin, P., Kools, P., Sciot, R., De Wever, I., Van Damme, B., Van de Ven, W. and Van den Berghe, H. (1993). Cytogenetic and fluorescence in situ hybridization investigation of ring chromosomes characterizing a specific pathologic subgroup of adipose tissue tumors. Cancer Genet Cytogenet **68**(2): 85-90.
Dang, C. V. (1999). c-Myc target genes involved in cell growth, apoptosis, and metabolism. Mol Cell Biol **19**(1): 1-11.
de Crombrugghe, B., Lefebvre, V., Behringer, R. R., Bi, W., Murakami, S. and Huang, W. (2000). Transcriptional mechanisms of chondrocyte differentiation. Matrix Biol **19**(5): 389-94.
de Crombrugghe, B., Lefebvre, V. and Nakashima, K. (2001). Regulatory mechanisms in the pathways of cartilage and bone formation. Curr Opin Cell Biol **13**(6): 721-7.
Dornan, J., Taylor, P. and Walkinshaw, M. D. (2003). Structures of immunophilins and their ligand complexes. Curr Top Med Chem **3**(12): 1392-409.
Engler, M. J., Richardson, C.C (1982). The enzymes, Academic Press, San Diego.
Enzinger, F. M., Weiss, S.W (1995). Soft Tissue Tumors, Mosby: S^{t} Louis.
Etzkorn, F. A., Chang, Z. Y., Stolz, L. A. and Walsh, C. T. (1994). Cyclophilin residues that affect noncompetitive inhibition of the protein serine phosphatase activity of calcineurin by the cyclophilin.cyclosporin A complex. Biochemistry **33**(9): 2380-8.
Fakata, K. L., Elmquist, W. F., Swanson, S. A., Vorce, R. L., Prince, C. and Stemmer, P. M. (1998). Cyclosporin A has low potency as a calcineurin inhibitor in cells expressing high levels of P-glycoprotein. Life Sci **62**(26): 2441-8.
Faragher, R. G. and Kipling, D. (1998). How might replicative senescence contribute to human ageing? Bioessays **20**(12): 985-91.
Ferreira, P. A., Nakayama, T. A., Pak, W. L. and Travis, G. H. (1996). Cyclophilin-related protein RanBP2 acts as chaperone for red/green opsin. Nature **383**(6601): 637-40.
Fesik, S. W., Gampe, R. T., Jr., Holzman, T. F., Egan, D. A., Edalji, R., Luly, J. R., Simmer, R., Helfrich, R., Kishore, V. and Rich, D. H. (1990). Isotope-edited NMR of cyclosporin A bound to cyclophilin: evidence for a trans 9,10 amide bond. Science **250**(4986): 1406-9.
Fischer, G., Bang, H. and Mech, C. (1984). [Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides]. Biomed Biochim Acta **43**(10): 1101-11.
Fischer, G. and Schmid, F. X. (1990). The mechanism of protein folding. Implications of in vitro refolding models for de novo protein folding and translocation in the cell. Biochemistry **29**(9): 2205-12.
Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T. and Schmid, F. X. (1989). Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins. Nature **337**(6206): 476-8.
Ford, H. L. and Pardee, A. B. (1999). Cancer and the cell cycle. J Cell Biochem **Suppl 32-33**: 166-72.
Forus, A., Berner, J. M., Meza-Zepeda, L. A., Saeter, G., Mischke, D., Fodstad, O. and Myklebost, O. (1998). Molecular characterization of a novel amplicon at 1q21-q22 frequently observed in human sarcomas. Br J Cancer **78**(4): 495-503.
Forus, A., Bjerkehagen, B., Sirvent, N., Meza-Zepeda, L. A., Coindre, J. M., Berner, J. M., Myklebost, O. and Pedeutour, F. (2001a). A well-differentiated liposarcoma with a new type of chromosome 12-derived markers. Cancer Genet Cytogenet **131**(1): 13-8.
Forus, A., Florenes, V. A., Maelandsmo, G. M., Meltzer, P. S., Fodstad, O. and Myklebost, O. (1993). Mapping of amplification units in the q13-14 region of chromosome 12 in human sarcomas: some amplica do not include MDM2. Cell Growth Differ **4**(12): 1065-70.
Forus, A., Larramendy, M. L., Meza-Zepeda, L. A., Bjerkehagen, B., Godager, L. H., Dahlberg, A. B., Saeter, G., Knuutila, S. and Myklebost, O. (2001b). Dedifferentiation of a well-differentiated liposarcoma to a highly malignant metastatic osteosarcoma: amplification of 12q14 at all stages and gain of 1q22-q24 associated with metastases. Cancer Genet Cytogenet **125**(2): 100-11.
Forus, A., Weghuis, D. O., Smeets, D., Fodstad, O., Myklebost, O. and Geurts van Kessel, A. (1995a). Comparative genomic hybridization analysis of human sarcomas: II. Identification of novel amplicons at 6p and 17p in osteosarcomas. Genes, Chromosomes & Cancer **14**(1): 15-21.

Forus, A., Weghuis, D. O., Smeets, D., Fodstad, O., Myklebost, O. and van Kessel, A. G. (1995b). Comparative genomic hybridization analysis of human sarcomas: I. Occurrence of genomic imbalances and identification of a novel major amplicon at 1q21-q22 in soft tissue sarcomas. Genes Chromosomes Cancer **14**(1): 8-14.

Fox, D. S. and Heitman, J. (2002). Good fungi gone bad: the corruption of calcineurin. Bioessays **24**(10): 894-903.

Franke, E. K. and Luban, J. (1996). Inhibition of HIV-1 replication by cyclosporine A or related compounds correlates with the ability to disrupt the Gag-cyclophilin A interaction. Virology **222**(1): 279-82.

Friedman, J. and Weissman, I. (1991). Two cytoplasmic candidates for immunophilin action are revealed by affinity for a new cyclophilin: one in the presence and one in the absence of CsA. Cell **66**(4): 799-806.

Fujimori, F., Gunji, W., Kikuchi, J., Mogi, T., Ohashi, Y., Makino, T., Oyama, A., Okuhara, K., Uchida, T. and Murakami, Y. (2001). Crosstalk of prolyl isomerases, Pin1/Ess1, and cyclophilin A. Biochem Biophys Res Commun **289**(1): 181-90.

Galat, A. (2003). Peptidylprolyl cis/trans isomerases (immunophilins): biological diversity--targets--functions. Curr Top Med Chem **3**(12): 1315-47.

Gershoni, J. M. and Palade, G. E. (1982). Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to a positively charged membrane filter. Anal Biochem **124**(2): 396-405.

Haendler, B. and Hofer, E. (1990). Characterization of the human cyclophilin gene and of related processed pseudogenes. Eur J Biochem **190**(3): 477-82.

Harlow, E., Lane, D. (1988). Using Antibodies A laboratory manual, Cold Spring Harbor Laboratory Press.

Hartwell, L., Weinert, T., Kadyk, L. and Garvik, B. (1994). Cell cycle checkpoints, genomic integrity, and cancer. Cold Spring Harb Symp Quant Biol **59**: 259-63.

Hartwell, L. H. (1991). Twenty-five years of cell cycle genetics. Genetics **129**(4): 975-80.

Helman, L. J. and Meltzer, P. (2003). Mechanisms of sarcoma development. Nat Rev Cancer **3**(9): 685-94.

Hesketh, R. E. (1997). The Oncogene and Tumor Suppressor Gene Facts Book, Academic Press.

Hirai, M., Yoshida, S., Kashiwagi, H., Kawamura, T., Ishikawa, T., Kaneko, M., Ohkawa, H., Nakagawara, A., Miwa, M. and Uchida, K. (1999). 1q23 gain is associated with progressive neuroblastoma resistant to aggressive treatment. Genes Chromosomes Cancer **25**(3): 261-9.

Hong, F., Lee, J., Song, J. W., Lee, S. J., Ahn, H., Cho, J. J., Ha, J. and Kim, S. S. (2002). Cyclosporin A blocks muscle differentiation by inducing oxidative stress and inhibiting the peptidyl-prolyl-cis-trans isomerase activity of cyclophilin A: cyclophilin A protects myoblasts from cyclosporin A-induced cytotoxicity. Faseb J **16**(12): 1633-5.

Horowitz, D. S., Lee, E. J., Mabon, S. A. and Misteli, T. (2002). A cyclophilin functions in pre-mRNA splicing. Embo J **21**(3): 470-80.

Howard, B. R., Vajdos, F. F., Li, S., Sundquist, W. I. and Hill, C. P. (2003). Structural insights into the catalytic mechanism of cyclophilin A. Nat Struct Biol **10**(6): 475-81.

Hsu, T., McRackan, D., Vincent, T. S. and Gert de Couet, H. (2001). Drosophila Pin1 prolyl isomerase Dodo is a MAP kinase signal responder during oogenesis. Nat Cell Biol **3**(6): 538-43.

Hu, C. Y., Allen, M. and Gyllensten, U. (1992). Effect of freezing of the PCR buffer on the amplification specificity: allelic exclusion and preferential amplification of contaminating molecules. PCR Methods Appl **2**(2): 182-3.

Hur, S. and Bruice, T. C. (2002). The mechanism of cis-trans isomerization of prolyl peptides by cyclophilin. J Am Chem Soc **124**(25): 7303-13.

Ingelfinger, D., Gothel, S. F., Marahiel, M. A., Reidt, U., Ficner, R., Luhrmann, R. and Achsel, T. (2003). Two protein-protein interaction sites on the spliceosome-associated human cyclophilin CypH. Nucleic Acids Res **31**(16): 4791-6.

Jarvik, J. W. and Telmer, C. A. (1998). Epitope tagging. Annu Rev Genet **32**: 601-18.

Jarvis, D. L., Fleming, J. A., Kovacs, G. R., Summers, M. D. and Guarino, L. A. (1990). Use of early baculovirus promoters for continuous expression and efficient processing of foreign gene products in stably transformed lepidopteran cells. Biotechnology (N Y) **8**(10): 950-5.

Jin, L. and Harrison, S. C. (2002). Crystal structure of human calcineurin complexed with cyclosporin A and human cyclophilin. Proc Natl Acad Sci U S A **99**(21): 13522-6.

Kaufman, R. J., Brown, P. C. and Schimke, R. T. (1979). Amplified dihydrofolate reductase genes in unstably methotrexate-resistant cells are associated with double minute chromosomes. Proc Natl Acad Sci U S A **76**(11): 5669-73.

Kawazu, K., Yamada, K., Nakamura, M. and Ota, A. (1999). Characterization of cyclosporin A transport in cultured rabbit corneal epithelial cells: P-glycoprotein transport activity and binding to cyclophilin. Invest Ophthalmol Vis Sci **40**(8): 1738-44.

Kielberg, V., Brünner, N., Briand, P. (1993). Celledyrkning: En praktisk håndbok i dyrkning af mammale celler, Foreningen av danske lægestudenters forlag: København.

King, A. A., Debaun, M. R., Riccardi, V. M. and Gutmann, D. H. (2000). Malignant peripheral nerve sheath tumors in neurofibromatosis 1. Am J Med Genet **93**(5): 388-92.
Kinzler, K. W. and Vogelstein, B. (1996). Lessons from hereditary colorectal cancer. Cell **87**(2): 159-70.
Kipling, D. and Faragher, R. G. (1999). Telomeres. Ageing hard or hardly ageing? Nature **398**(6724): 191, 193.
Kistner, A., Gossen, M., Zimmermann, F., Jerecic, J., Ullmer, C., Lubbert, H. and Bujard, H. (1996). Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. Proc Natl Acad Sci U S A **93**(20): 10933-8.
Klee, C. B., Crouch, T. H. and Krinks, M. H. (1979). Calcineurin: a calcium- and calmodulin-binding protein of the nervous system. Proc Natl Acad Sci U S A **76**(12): 6270-3.
Klein, G. (1998). Foulds' dangerous idea revisited: the multistep development of tumors 40 years later. Adv Cancer Res **72**: 1-23.
Knudson, A. G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A **68**(4): 820-3.
Kohno, T. and Yokota, J. (2002). [Genetic factors involved in cancer susceptibility]. Gan To Kagaku Ryoho **29**(7): 1255-62.
Kopito, R. R. and Sitia, R. (2000). Aggresomes and Russell bodies. Symptoms of cellular indigestion? EMBO Rep **1**(3): 225-31.
Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res **15**(20): 8125-48.
Kudoh, K., Takano, M., Koshikawa, T., Hirai, M., Yoshida, S., Mano, Y., Yamamoto, K., Ishii, K., Kita, T., Kikuchi, Y., Nagata, I., Miwa, M. and Uchida, K. (1999). Gains of 1q21-q22 and 13q12-q14 are potential indicators for resistance to cisplatin-based chemotherapy in ovarian cancer patients. Clin Cancer Res **5**(9): 2526-31.
Lee, S. P., Hwang, Y. S., Kim, Y. J., Kwon, K. S., Kim, H. J., Kim, K. and Chae, H. Z. (2001). Cyclophilin a binds to peroxiredoxins and activates its peroxidase activity. J Biol Chem **276**(32): 29826-32.
Lefebvre, V., Behringer, R. R. and de Crombrugghe, B. (2001). L-Sox5, Sox6 and Sox9 control essential steps of the chondrocyte differentiation pathway. Osteoarthritis Cartilage **9 Suppl A**: S69-75.
Lefebvre, V., Li, P. and de Crombrugghe, B. (1998). A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. Embo J **17**(19): 5718-33.
Lengauer, C., Kinzler, K. W. and Vogelstein, B. (1998). Genetic instabilities in human cancers. Nature **396**(6712): 643-9.
Leverson, J. D. and Ness, S. A. (1998). Point mutations in v-Myb disrupt a cyclophilin-catalyzed negative regulatory mechanism. Mol Cell **1**(2): 203-11.
Li, F. P., Fraumeni, J. F., Jr., Mantel, N. and Miller, R. W. (1969). Cancer mortality among chemists. J Natl Cancer Inst **43**(5): 1159-64.
Li, W. and Handschumacher, R. E. (2002). Identification of two calcineurin B-binding proteins: tubulin and heat shock protein 60. Biochim Biophys Acta **1599**(1-2): 72-81.
Lichtman, J. W. (1994). Confocal microscopy.
Lindahl, T. E. (1996). Genetic instability in cancer, Cold Spring Harbor Press.
Lodish, B., Harvey et al. (2000). Molecular Cell Biology, W.H.Freeeman and co.
Lygren, B. (2002). Characterization of *COAS1* and *COAS2*, two novel genes in 1q21 frequently amplified and over-expressed in sarcomas. Department of Biochemistry/Department of Tumor Biology. Oslo, University of Oslo**:** 132.
Maelandsmo, G. M., Berner, J. M., Florenes, V. A., Forus, A., Hovig, E., Fodstad, O. and Myklebost, O. (1995). Homozygous deletion frequency and expression levels of the CDKN2 gene in human sarcomas--relationship to amplification and mRNA levels of CDK4 and CCND1. Br J Cancer **72**(2): 393-8.
Maelandsmo, G. M., Florenes, V. A., Hovig, E., Oyjord, T., Engebraaten, O., Holm, R., Borresen, A. L. and Fodstad, O. (1996). Involvement of the pRb/p16/cdk4/cyclin D1 pathway in the tumorigenesis of sporadic malignant melanomas. Br J Cancer **73**(8): 909-16.
Malkin, D., Li, F. P., Strong, L. C., Fraumeni, J. F., Jr., Nelson, C. E., Kim, D. H., Kassel, J., Gryka, M. A., Bischoff, F. Z., Tainsky, M. A. and et al. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science **250**(4985): 1233-8.
Malumbres, M., Perez De Castro, I., Hernandez, M. I., Jimenez, M., Corral, T. and Pellicer, A. (2000). Cellular Response to Oncogenic Ras Involves Induction of the Cdk4 and Cdk6 Inhibitor p15INK4b. Mol. Cell. Biol. **20**(8): 2915-2925.
Matsuda, S. and Koyasu, S. (2003). Regulation of MAPK signaling pathways through immunophilin-ligand complex. Curr Top Med Chem **3**(12): 1358-67.

Maxwell, P. H., Dachs, G. U., Gleadle, J. M., Nicholls, L. G., Harris, A. L., Stratford, I. J., Hankinson, O., Pugh, C. W. and Ratcliffe, P. J. (1997). Hypoxia-inducible factor-1 modulates gene expression in solid tumors and influences both angiogenesis and tumor growth. Proc Natl Acad Sci U S A **94**(15): 8104-9.
McAlpine, P. (1995). Genetic nomenclature guide. Human. Trends Genet: 39-42.
Meltzer, P. S. (2001a). Large Scale genome analysis. Bioinformatics; A Pratical Guide to the Analysis of Genes and Proteins. A. D. Baxevanis and B. F. Francis Ovellette, Wiley-Interscience.
Meltzer, P. S. (2001b). Large-scale genome analysis. Methods Biochem Anal **43**: 393-412.
Mensa-Wilmot, K. and Englund, P. T. (1992). Non-fusion proteins expressed in E. coli: detection by alpha-complementation (blue/white selection). Nucleic Acids Res **20**(1): 143.
Messenger, M. M., Saulnier, R. B., Gilchrist, A. D., Diamond, P., Gorbsky, G. J. and Litchfield, D. W. (2002). Interactions between protein kinase CK2 and Pin1. Evidence for phosphorylation-dependent interactions. J Biol Chem **277**(25): 23054-64.
Meza-Zepeda, L. A., Berner, J. M., Henriksen, J., South, A. P., Pedeutour, F., Dahlberg, A. B., Godager, L. H., Nizetic, D., Forus, A. and Myklebost, O. (2001). Ectopic sequences from truncated HMGIC in liposarcomas are derived from various amplified chromosomal regions. Genes Chromosomes Cancer **31**(3): 264-73.
Meza-Zepeda, L. A., Forus, A., Lygren, B., Dahlberg, A. B., Godager, L. H., South, A. P., Marenholz, I., Lioumi, M., Florenes, V. A., Maelandsmo, G. M., Serra, M., Mischke, D., Nizetic, D., Ragoussis, J., Tarkkanen, M., Nesland, J. M., Knuutila, S. and Myklebost, O. (2002). Positional cloning identifies a novel cyclophilin as a candidate amplified oncogene in 1q21. Oncogene **21**(14): 2261-9.
Michnick, S. W., Rosen, M. K., Wandless, T. J., Karplus, M. and Schreiber, S. L. (1991). Solution structure of FKBP, a rotamase enzyme and receptor for FK506 and rapamycin. Science **252**(5007): 836-9.
Miller, L. D., Long, P. M., Wong, L., Mukherjee, S., McShane, L. M. and Liu, E. T. (2002). Optimal gene expression analysis by microarrays. Cancer Cell **2**(5): 353-61.
Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb Symp Quant Biol **51 Pt 1**: 263-73.
Myklebost, O. (1998). [Molecular biology diagnosis of sarcoma]. Tidsskr Nor Laegeforen **118**(14): 2188-93.
Neal, J. W. and Clipstone, N. A. (2002). Calcineurin mediates the calcium-dependent inhibition of adipocyte differentiation in 3T3-L1 cells. J Biol Chem **277**(51): 49776-81.
Nevins, J. R. (2001). The Rb/E2F pathway and cancer. Hum Mol Genet **10**(7): 699-703.
Nilsen, T., Slagsvold, T., Skjerpen, C. S., Brech, A., Stenmark, H. and Olsnes, S. (2004). Peroxisomal targeting as a tool for assaying potein-protein interactions in the living cell: cytokine-independent survival kinase (CISK) binds PDK-1 in vivo in a phosphorylation-dependent manner. J Biol Chem **279**(6): 4794-801.
Nilsson, M., Meza-Zepeda, L. A., Mertens, F., Forus, A., Myklebost, O. and Mandahl, N. (2004). Amplification of chromosome 1 sequences in lipomatous tumors and other sarcomas. Int J Cancer **109**(3): 363-9.
Nishida, T., Hirota, S., Taniguchi, M., Hashimoto, K., Isozaki, K., Nakamura, H., Kanakura, Y., Tanaka, T., Takabayashi, A., Matsuda, H. and Kitamura, Y. (1998). Familial gastrointestinal stromal tumours with germline mutation of the KIT gene. Nat Genet **19**(4): 323-4.
O'Gorman, S., Fox, D. T. and Wahl, G. M. (1991). Recombinase-mediated gene activation and site-specific integration in mammalian cells. Science **251**(4999): 1351-5.
Pardal, R., Clarke, M. F. and Morrison, S. J. (2003). Applying the principles of stem-cell biology to cancer. Nat Rev Cancer **3**(12): 895-902.
Pardee, A. B. (1994). Growth dysregulation in cancer cells. Adv Cancer Res **65**: 213-28.
Pardee, A. B. (2003). Normal cells arrest cancer cells in G2. Cell Cycle **2**(5): 454.
Pemberton, T. J. and Kay, J. E. (2003). Cyclophilin sensitivity to sanglifehrin A can be correlated to the same specific tryptophan residue as cyclosporin A. FEBS Lett **555**(2): 335-40.
Picken, N. C., Eschenlauer, S., Taylor, P., Page, A. P. and Walkinshaw, M. D. (2002). Structural and biological characterisation of the gut-associated cyclophilin B isoforms from Caenorhabditis elegans. J Mol Biol **322**(1): 15-25.
Pinson, B., Brendeford, E. M., Gabrielsen, O. S. and Daignan-Fornier, B. (2001). Highly conserved features of DNA binding between two divergent members of the Myb family of transcription factors. Nucleic Acids Res **29**(2): 527-35.
Quackenbush, J. (2002). Microarray data normalization and transformation. Nat Genet **32 Suppl**: 496-501.
Rabbitts, P. H. (1994). Genetic changes in the development of lung cancer. Br Med Bull **50**(3): 688-97.
Ratajczak, T., Ward, B. K. and Minchin, R. F. (2003). Immunophilin chaperones in steroid receptor signalling. Curr Top Med Chem **3**(12): 1348-57.
Read, S. a. (1999). Human Molecular Genetics 2. Oxford, BIOS Scientific Publishers, Ltd.
Reidt, U., Wahl, M. C., Fasshauer, D., Horowitz, D. S., Luhrmann, R. and Ficner, R. (2003). Crystal structure of a complex between human spliceosomal cyclophilin H and a U4/U6 snRNP-60K peptide. J Mol Biol **331**(1): 45-56.

Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M. and Fischer, G. (1998). Side-chain effects on peptidyl-prolyl cis/trans isomerisation. J Mol Biol **279**(2): 449-60.
Renart, J. and Sandoval, I. V. (1984). Western blots. Methods Enzymol **104**: 455-60.
Rous, P. (1911). J Exp Med **13**: 397.
Ryo, A., Nakamura, M., Wulf, G., Liou, Y. C. and Lu, K. P. (2001). Pin1 regulates turnover and subcellular localization of beta-catenin by inhibiting its interaction with APC. Nat Cell Biol **3**(9): 793-801.
Saeter, G., Elomaa, I., Wahlqvist, Y., Alvegard, T. A., Wiebe, T., Monge, O., Forrestier, E. and Solheim, O. P. (1997). Prognostic factors in bone sarcomas. Acta Orthop Scand Suppl **273**: 156-60.
Sambrook, J., Russell D.W. (2001). Molecular Cloning, A laboratory manual, Cold Spring Harbor Laboratory Press.
Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A **74**(12): 5463-7.
Saphire, A. C., Bobardt, M. D. and Gallay, P. A. (1999). Host cyclophilin A mediates HIV-1 attachment to target cells via heparans. Embo J **18**(23): 6771-85.
Sauer, B. (1994). Site-specific recombination: developments and applications. Curr Opin Biotechnol **5**(5): 521-7.
Schena, M., Shalon, D., Davis, R. W. and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science **270**(5235): 467-70.
Schiene-Fischer, C. and Yu, C. (2001). Receptor accessory folding helper enzymes: the functional role of peptidyl prolyl cis/trans isomerases. FEBS Letters **495**(1-2 SU -): 1-6.
Schubert, A. and Grimm, S. (2004). Cyclophilin D, a component of the permeability transition-pore, is an apoptosis repressor. Cancer Res **64**(1): 85-93.
Schwab, M. (1999). Oncogene amplification in solid tumors. Semin Cancer Biol **9**(4): 319-25.
Sehgal, S. N., Baker, H. and Vezina, C. (1975). Rapamycin (AY-22,989), a new antifungal antibiotic. II. Fermentation, isolation and characterization. J Antibiot (Tokyo) **28**(10): 727-32.
Shaw, P. E. (2002). Peptidyl-prolyl isomerases: a new twist to transcription. EMBO Rep **3**(6): 521-6.
Sherry, B., Zybarth, G., Alfano, M., Dubrovsky, L., Mitchell, R., Rich, D., Ulrich, P., Bucala, R., Cerami, A. and Bukrinsky, M. (1998). Role of cyclophilin A in the uptake of HIV-1 by macrophages and T lymphocytes. Proc Natl Acad Sci U S A **95**(4): 1758-63.
Simonsen, J. L., Rosada, C., Serakinci, N., Justesen, J., Stenderup, K., Rattan, S. I., Jensen, T. G. and Kassem, M. (2002). Telomerase expression extends the proliferative life-span and maintains the osteogenic potential of human bone marrow stromal cells. Nat Biotechnol **20**(6): 592-6.
Simpson, R. J. (2003). Proteins and Proteomics - A laboratory manual, Cold Spring Harbor Laboratory Press.
Skehan, P., Storeng, R., Scudiero, D., Monks, A., McMahon, J., Vistica, D., Warren, J. T., Bokesch, H., Kenney, S. and Boyd, M. R. (1990). New colorimetric cytotoxicity assay for anticancer-drug screening. J Natl Cancer Inst **82**(13): 1107-12.
Slominski, A., Wortsman, J., Carlson, A., Mihm, M., Nickoloff, B. and McClatchey, K. (1999). Molecular pathology of soft tissue and bone tumors. A review. Arch Pathol Lab Med **123**(12): 1246-59.
Smith, T. R., Levine, E. A., Perrier, N. D., Miller, M. S., Freimanis, R. I., Lohman, K., Case, L. D., Xu, J., Mohrenweiser, H. W. and Hu, J. J. (2003). DNA-repair genetic polymorphisms and breast cancer risk. Cancer Epidemiol Biomarkers Prev **12**(11 Pt 1): 1200-4.
Smits, P., Li, P., Mandel, J., Zhang, Z., Deng, J. M., Behringer, R. R., de Crombrugghe, B. and Lefebvre, V. (2001). The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. Dev Cell **1**(2): 277-90.
Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol **98**(3): 503-17.
Steen, H. B. (1991). Flow cytometry instrumentation. Particle Analysis in Oceanography. S. Demers. Heidelberg, Springer-Verlag. **G 27**.
Szymanska, K. and Hainaut, P. (2003). TP53 and mutations in human cancer. Acta Biochim Pol **50**(1): 231-8.
Sæter, G. (1996). Treatment of bone and soft tissue sarcomas. The role of chemotherapy, dose-response relationships and haematopoietic growth factors, Gardiner-Caldwell Communications, Chesire.
Saal, L. H., Troein, C., Vallon-Christersson, J., Gruvberger, S., Borg, A. and Peterson, C. (2002). BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. Genome Biol **3**(8): SOFTWARE0003.
Takahashi, N., Hayano, T. and Suzuki, M. (1989). Peptidyl-prolyl cis-trans isomerase is the cyclosporin A-binding protein cyclophilin. Nature **337**(6206): 473-5.
Tarkkanen, M., Huuhtanen, R., Virolainen, M., Wiklund, T., Asko-Seljavaara, S., Tukiainen, E., Lepantalo, M., Elomaa, I. and Knuutila, S. (1999). Comparison of genetic changes in primary sarcomas and their pulmonary metastases. Genes Chromosomes Cancer **25**(4): 323-31.
Tatusova, T. A. and Madden, T. L. (1999). BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. FEMS Microbiol Lett **174**(2): 247-50.

Tiberghien, F., Wenandy, T. and Loor, F. (2000). The potent immunosuppressive cyclosporin FR901459 inhibits the human P-glycoprotein and formyl peptide receptor functions. J Antibiot (Tokyo) **53**(5): 509-15.
Tsien, R. Y. (1998). The green fluorescent protein. Annu Rev Biochem **67**: 509-44.
Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. and Clardy, J. (1991). Atomic structure of FKBP-FK506, an immunophilin-immunosuppressant complex. Science **252**(5007): 839-42.
Varmus, H. E., Padgett, T., Heasley, S., Simon, G. and Bishop, J. M. (1977). Cellular functions are required for the synthesis and integration of avian sarcoma virus-specific DNA. Cell **11**(2): 307-19.
Varmus, H. E., Weinberg, R.A. (1993). Genes and the biology of cancer, Scientific American Library.
Vogelstein, B. and Kinzler, K. W. (1993). The multistep nature of cancer. Trends Genet **9**(4): 138-41.
Walsh, C. T., Zydowsky, L. D. and McKeon, F. D. (1992). Cyclosporin A, the cyclophilin class of peptidylprolyl isomerases, and blockade of T cell signal transduction. J Biol Chem **267**(19): 13115-8.
Weinberg, R. A. (1996). How cancer arises. Sci Am **275**(3): 62-70.
Wickham, T. J., Davis, T., Granados, R. R., Shuler, M. L. and Wood, H. A. (1992). Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system. Biotechnol Prog **8**(5): 391-6.
Wigler, P. W. (1999). PSC833, cyclosporin A, and dexniguldipine effects on cellular calcein retention and inhibition of the multidrug resistance pump in human leukemic lymphoblasts. Biochem Biophys Res Commun **257**(2): 410-3.
Wilfinger, W. W., Mackey, K. and Chomczynski, P. (1997). Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. Biotechniques **22**(3): 474-6, 478-81.
Willenbrink, W., Halaschek, J., Schuffenhauer, S., Kunz, J. and Steinkasserer, A. (1995). Cyclophilin A, the major intracellular receptor for the immunosuppressant cyclosporin A, maps to chromosome 7p11.2-p13: four pseudogenes map to chromosomes 3, 10, 14, and 18. Genomics **28**(1): 101-4.
Wulf, G. M., Liou, Y. C., Ryo, A., Lee, S. W. and Lu, K. P. (2002). Role of Pin1 in the regulation of p53 stability and p21 transactivation, and cell cycle checkpoints in response to DNA damage. J Biol Chem **277**(50): 47976-9.
Yang, W. M., Inouye, C. J. and Seto, E. (1995). Cyclophilin A and FKBP12 interact with YY1 and alter its transcriptional activity. J Biol Chem **270**(25): 15187-93.
Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J. and Speed, T. P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res **30**(4): e15.
Yao, F., Svensjo, T., Winkler, T., Lu, M., Eriksson, C. and Eriksson, E. (1998). Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. Hum Gene Ther **9**(13): 1939-50.
Yu, J., Zhang, L., Hwang, P. M., Rago, C., Kinzler, K. W. and Vogelstein, B. (1999). Identification and classification of p53-regulated genes. Proc Natl Acad Sci U S A **96**(25): 14517-22.
Yu, Q., Geng, Y. and Sicinski, P. (2001). Specific protection against breast cancers by cyclin D1 ablation. Nature **411**(6841): 1017-21.
Yurchenko, V., Zybarth, G., O'Connor, M., Dai, W. W., Franchin, G., Hao, T., Guo, H., Hung, H. C., Toole, B., Gallay, P., Sherry, B. and Bukrinsky, M. (2002). Active site residues of cyclophilin A are crucial for its signaling activity via CD147. J Biol Chem **277**(25): 22959-65.
Zakian, V. A. (1997). Life and cancer without telomerase. Cell **91**(1): 1-3.
Zhang, Z., Carriero, N. and Gerstein, M. (2004). Comparative analysis of processed pseudogenes in the mouse and human genomes. Trends Genet **20**(2): 62-7.
Zhao, Q., Eberspaecher, H., Lefebvre, V. and De Crombrugghe, B. (1997). Parallel expression of Sox9 and Col2a1 in cells undergoing chondrogenesis. Dev Dyn **209**(4): 377-86.

Appendix A: Glossary

actin An abundant 43-kd protein that polymerizes to form cytoskeletal filaments.

active site The region of an enzyme that binds substrates and catalyses an enzymatic reaction.

allele One copy of a gene.

amino acid Monomeric building blocks of proteins, consisting of a carbon atom bound to a carboxyl group, an amino group, a hydrogen atom, and a distinctive side chain.

amplicon A chromosomal region amplified in certain subtypes of cancer

anaphase The phase of mitosis during which sister chromatids separate and move to opposite poles of the spindle.

APC; anaphase-promoting complex A ubiquitin ligase that triggers progression from metaphase to anaphase by signaling the degradation of cyclin B and cohesins.

anaplastic A term used to describe cancer cells that divide rapidly and have little or no resemblance to normal cells.

angiogenesis The formation of new blood vessels.

antibody A protein produced by B lymphocytes that binds to an antigen, (foreign molecule), via its *epitope.*

antigen A molecule against which an antibody is directed.

apoptosis An active process of programmed cell death, characterised by cleavage of chromosomal DNA, chromatin condensation, and fragmentation of both the nucleus and the cell.

autocrine signaling A type of cell signaling in which a cell produces a growth factor to which it also responds.

autoradiography The detection of radioisotopically labeled molecules by exposure to X-ray film.

bacteriophage A bacterial virus.

base-excision repair A mechanism of DNA repair in which single damaged bases are removed from a DNA molecule.

benign tumour A tumour that remains confined to its site of origin.

blastula an early embryonic form produced by cleavage of a fertilised ovum and usually consisting of a single layer of cells surrounding a fluid-filled spherical cavity.

calmodulin A calcium-binding protein.

cancer stem cell A cancerous cell with either aquired ability of self-renewal, or a mutated stem cell acting as a cancer cell precursor. Thought to mediate growth with or without the aid of more differentiated cells in a tumour. Identified per January 2004 in glioma and leukaemia.

capillary electrophoresis The separation of biomolecules in a buffer-filled narrow capillary by application of an electric field, such as in isoelectric focusing, IEF. (Separation by electrophoresis is based on solute velocity differences in an electric field. The velocity of an ion is a function of its electrophoretic mobility and the applied voltage.)

carcinogen A cancer-inducing agent.

carcinoma A cancer that arises in epithelial cells, of endo- or ectodermal origin.

cardiac hypertrophy Enlargement of either one or both heart ventricles. This increase in ventricular mass is attributed to sustained abnormal pressure or volume loads and is a contributor to cardiovascular morbidity and mortality. Common cause of death in industrialized countries worldwide.

Cdk inhibitor (CKI) A family of proteins that bind Cdks and inhibit their activity.

Cdks Cyclin dependent protein kinases that control the cell cycle of eukaryotes.

cell cycle Regulated series of events occurring in the eukaryotic cell between cell divisions.

chaotropic Pertaining to *chaotropism,* often ions with chaotropic ability.

chaotropism (Greek: Chaos = disorder, confusion; trope = a turning) The property of certain substances, usually ions (e.g., SCN^-, ClO^{4-}, guanidinium), to disrupt the structure of water and thereby promote the solubility of nonpolar substances in polar solvents (e.g. water), the unfolding of proteins, the elution from or movement through a chromatographic medium of an otherwise tightly bound substance, etc.

chaperone A protein that facilitates the correct folding or assembly of other proteins.

chaperonin A family of heat-shock proteins that assist during protein folding by keeping amino acid chains outstretched until proper folding is assisted by *chaperones*.

clonogeneicity The ability of a single cell to form clonal expansions, thus grow to form tumours *in vitro*.

collagen The major structural protein of the extracellular matrix.

complementary DNA (cDNA) A DNA molecule that is complementary to a mRNA molecule, synthesized *in vitro* by reverse transcriptase.

cosmid A vector that contains bacteriophage □ sequences, antibiotic resistance sequences, and an origin of replication. It can accommodate large DNA inserts of up to 45 kb.

cyclins A family of proteins that regulate the activity of Cdks and control progression through the cell cycle.

cyclophilins A protein family with peptidyl-prolyl-isomerase activity, involved in protein folding and diverse regulatory processes. Belongs to the immunophilin superfamily of proteins.

double minutes Characteristic chromosomal arrangement of amplified regions in both murine and human tumours.

ectoderm The outer germ layer; gives rise to tissues that include the skin and nervous system.

embryogenesis Generation of an embryo from a fertilised egg.

endoderm The inner germ layer; gives rise to internal organs.

endoplasmic reticulum (ER) An extensive network of membrane-enclosed tubules and sacs involved in protein sorting and processing as well as in lipid synthesis.

enzymes Proteins or RNAs that catalyze biological reactions.

erythrocytes Red blood cells.

exon A segment of a gene that contains a coding sequence.

exonuclease An enzyme that hydrolyzes DNA molecules in either the 5′ to 3′ or 3′ to 5′ direction, beginning at the ends of the linear molecules in contrast to *endonucleases*.

expression vector A vector used to direct expression of a cloned DNA fragment in a host cell, e.g. *plasmid*.

extinction coefficient or molar absorption coefficient Unit describing the molar cross-section for absorption of light. The greater the cross-section of the absorbing molecule, the greater its ability to block the passage of the incident radiation. Its dimensions are 1/(concentration x length), and it is normally expressed in litres per mole per centimetre ($L\ mol^{-1}\ cm^{-1}$).

extracellular matrix Secreted proteins and polysaccharides that fill spaces between cells and bind cells and tissues together.

fibroblast A cell type found in connective tissue.

filamins Structural protein connecting actin filaments in eukaryotic cells, stabilizing the meshwork and greatly increasing the viscosity of the surrounding medium.

fluorophore A molecule that may be linked to a non-fluorescent molecule for detection purposes through fluorescence techniques.

G_0 A quiescent state in which cells remain metabolically active but do not proliferate.

G_1 phase The phase of the cell cycle between the end of mitosis and the beginning of DNA synthesis.

G_2 phase The phase of the cell cycle between the end of S phase and the beginning of mitosis.

gene A segment of DNA that encodes a polypeptide chain or a RNA molecule.

gene amplification An increase in the number of copies of a gene resulting from the increased replication of a region of DNA.

giant marker, ring and rod chromosomes Chromosomal rearrangements containing amplified regions characteristic of well-differentiated liposarcomas (WDLPS).

glycosaminoglycan (GAG) Glycosaminoglycans are polysaccharides (e.g. chondroitin sulfate, dermatan sulfate, heparan sulfate, heparin and keratan sulfate), made of repeating disaccharides (usually 40-100 times), which consist of uronic acid (or galactose) and hexosamines. See *proteoglycans*.

green fluorescent protein (GFP) A protein from the jellyfish *A. .victoria* that is commonly used as a marker for fluorescence microscopy.

growth factors Polypeptides that control animal cell growth and differentiation.

heat-shock proteins A highly conserved group of chaperone proteins expressed in cells exposed to elevated temperatures or other forms of environmental stress.

histocytes Or macrophages, cells with endocytic ability. Part of the immune system in vivo.

homeostasis The maintenance of a dynamic steady state by regulatory mechanisms that compensate for changes in external circumstances.

hyperlipidemia A general term for elevated concentrations of any or all of the lipids in the plasma, such as cholesterol, triglycerides and lipoproteins.

hyperplasia Too much growth of cells or tissue in a specific area, exemplified by increased growth in the lining of the prostate, causing benign prostatic hyperplasia.

hypoxia Metabolic condition in which the supply of oxygen is severely limited

***in situ* hybridization** The use of radioactive or fluorescent probes to detect RNA or DNA sequences in cell extracts, chromosomes, or intact cells.

intercalating dye A planar molecule that is inserted between two successive bases in a nucleic acid, such as ethidium bromide, EtBr.
***in vitro* translation** Protein synthesis in a cell-free extract. (Confer the RTS system described in this work.)

interphase The period of the cell cycle between mitoses, which includes G_1, S, and G_2 phases.

intracellular signal transduction A chain of reactions that transmits chemical signals from the cell surface to their intracellular targets, see the *SOS-RAS-RAF-MAP kinase* pathway.

intron Noncoding sequence that interrupts exons in a gene.

karyotype Chromosomal content of an eukaryotic cell, displayed in terms of its appearance at metaphase in a karyogram.

Lambert-Beer law log (I_0/I) = εcl Incident light intensity (I_0) divided by transmitted light intensity (I) equals the molar extinction coefficient (ε) in the medium times the concentration (c) of the absorbing species times the cuvette's path length (l). This law collects the basics for spectrophotometric quantification of biomolecules in a sample, where a fixed path length causes a proportional relationship between the absorbance A (= $\log(I_0/I)$) and the concentration of the absorbing solute c. Empirically deduced law.

leucine zipper A protein dimerization domain containing repeated leucine residues; found in many transcription factors, often as basic leucine zippers, bLH.

leukemia Cancer arising from the precursors of circulating blood cells.

liposome A small spherical vesicle composed of a phospholipid bilayer, which forms spontaneously when phospholipids are suspended in an aqueous buffer.

lymphocyte A blood cell that functions in the immune response. B lymphocytes produce *antibodies* and T lymphocytes are responsible for cell mediated immunity.

M phase The mitotic phase of the cell cycle.

malignant tumour A tumour that invades normal tissue and spreads throughout the body.

MAP kinases A family of mitogen-activated protein-serine/threonine kinases that are ubiquitous regulators of cell growth and differentiation.

matrix The inner mitochondrial space.

mesoderm The middle germ layer; gives rise to connective tissues and the hematopoietic system.

methylation Addition of methyl ($-CH_3$) groups to positions on DNA strands. Functions mainly as a negative regulator of transcription in eukaryotes, through methylation of CpG islands upstream of actively transcribed genes.

messenger RNA (mRNA) A RNA molecule that serves as a template for protein synthesis.

metaphase The phase of mitosis during which the chromosomes are aligned on a metaphase plate in the center of the cell.

metastasis Spread of cancer cells through the blood or lymphatic system to other organ sites.

mismatch repair Repair system that removes mismatched bases from newly synthesized DNA strands.

monoclonal antibody An antibody produced by a clonal line of B lymphocytes.

mutation Genetic alteration.

myoblast A skeletal muscle cell precursor.
myosin A protein that interacts with actin, functions as a molecular motor.

neoplasia Abnormal and uncontrolled cell growth

neurite extension Extension of a neuronal cell, the term neurite may signify both dendrite and axon.

nucleoside A purine or pyrimidine base linked to a sugar (ribose or deoxyribose).

nucleotide A phosphorylated nucleoside.

nucleotide excision repair A mechanism of DNA repair in which oligonucleotides containing damaged bases are removed from a DNA molecule.

oligonucleotide Short polymer of only a few nucleotides.

oncogene Gene capable of inducing one or more characteristics of cancer cells.

palindrome Segment of duplex DNA in which the base sequences of the two strands exhibit twofold rotational symmetry about an axis.

peptidyl prolyl isomerase An enzyme that facilitates protein folding by catalyzing the *cis-trans* isomerization of prolyl peptide bonds.

phalloidin A drug that binds to actin filaments and prevents their disassembly.

phosphorylation The addition of a phosphate group to a molecule.

plasmid A small, circular DNA molecule capable of independent replication in a host cell. See *vector*.

polylinker A short, often synthetic fragment of DNA containing recognition sequences for several restriction endonucleases

polymerase chain reaction (PCR) A method for amplifying a region of DNA by repeated cycles of DNA synthesis *in vitro*.

processed pseudogene A pseudogene that has arisen by reverse transcription of mRNA.

prokaryotic cells Cells lacking a nuclear envelope, cytoplasmic organelles, and a cytoskeleton (primarily bacteria).

promoter A DNA sequence to which RNA polymerase binds to initiate transcription.

proofreading The selective removal of mismatched bases by DNA polymerase.

proteasome A large 20S protease complex that degrades proteins tagged by *ubiquitin.*

protein kinase An enzyme that phosphorylates proteins by transferring a phosphate group from ATP, see also *phosphorylation.*

protein phosphatase An enzyme that reverses the action of protein kinases by removing phosphate groups from phosphorylated proteins.

proteoglycan A protein containing one or more covalently linked and usually sulfated glycosaminoglycan. Aggrecan, for example, is the major proteoglycan component in articular cartilage.

proteolysis Degradation of polypeptide chains.

proto-oncogene A normal cell gene that can be converted into an oncogene.

pseudogene A nonfunctional gene copy.

Raf A protein-serine/threonine kinase (encoded by the *RAF* oncogene) that is activated by Ras and leads to activation of MAP kinase.

Ras A family of small GTP binding proteins (encoded by the *RAS* oncogenes) that couple growth factor receptors to intracellular targets, including the Raf protein-serine/threonine kinase and the MAP kinase pathway.

RB A transcriptional regulatory protein encoded by the tumour suppressor gene *RB1* that was identified by the genetic analysis of hereditary retinoblastoma, a cancer that causes retinal tumours at young age.

recombinant molecule A DNA insert joined to a vector.

repressor A regulatory molecule that blocks transcription. Example: The Tet-repressor molecule expressed in the Flp-In system used in this work.

restriction endonuclease An enzyme that cleaves DNA internally at a specific sequence, often at *palindromic* repeats.

restriction point A regulatory point in animal cell cycles that occurs late in G_1. After this point, a cell is committed to entering S and undergoing one cell division cycle.

reverse transcriptase A DNA polymerase that uses an RNA template to generate complementary DNA, (cDNA).

ribosomal RNA (rRNA) The RNA component of ribosomes, in humans the 5S, 18S and 28S fractions together constitute the total rRNA.

ribosomes Particles composed of RNA and proteins that are the sites of protein synthesis.

Rotamase An immunophilin, enzyme with peptidylprolyl isomerase activity.

Rous sarcoma virus (RSV) An acutely transforming retrovirus, in which the first oncogene was identified.

S phase The phase of the cell cycle during which DNA replication occurs.

sarcoma Cancer of cells of connective tissue.

sarcosyl N-lauroyl-sarcosine, anionic detergent used for RNA isolation and RNA stabilization in solution.

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) A commonly used method to separate proteins by gel electrophoresis on the basis of size.

SH2 domain A protein domain of approximately 100 amino acids that binds phosphotyrosine-containing peptides. See the *SOS-RAS-RAF-MAP kinase pathway*.

signal sequence A hydrophobic sequence at the amino terminus of a polypeptide chain that targets it for secretion or organelle inclusion after modification in the endoplasmatic reticulum, (the latter for eukaryotic cells only).

smooth muscle Muscles surrounding internal organs. Contraction of smooth muscle controls the diameter of blood vessels and propels food along the gastrointestinal tract, amongst other functions.

SOS-RAS-RAF-MAP kinase Pathway signal cascade where ligand binding to a receptor is carried via several cytoplasmic proteins to SOS-RAS-RAF and ultimately forms a MAP kinase dimer that translocates to the nucleus where it activates transcription of cell cycle- and differentiation-specific proteins. Oncogenic activation of the Ras or both Ras and Raf protein(s) is common in many cancer types.

Src A nonreceptor protein-tyrosine kinase encoded by the oncogene (*src*) of Rous sarcoma virus.

stem cell A cell that divides to produce daughter cells that can either differentiate or remain as stem cells. Most stem cells divide asymmetrically to give one stem cell with self renewal potential equal to the parent cell, and one stem cell intended for differentiation.

striated muscle Cross striations is a characteristic trait in differentiated muscles associated with contractile ability of mature (skeletal, heart and smooth) muscle cells.

stromal cells Cells forming supporting connective tissue of an organ.

substrate A molecule acted upon by an enzyme.

supernumerary rings Chromosomal arrangement trait characteristic of well-differentiated liposarcomas (WDLPS). These rings contain amplified regions of DNA.

Svedberg (S) Unit of measure of the rate at which a particle sediments in a centrifugal field.

Synovial sarcoma Relatively rare type of soft tissue sarcoma most common at knee.

Synovial membrane A layer of connective tissue that lines the cavities of joints, tendon sheaths, bursae (fluid-filled sacs between tendons and bones), and the cavity (hollow enclosed area) that separates the bones of a freely movable joint, such as the knee or elbow. The synovial membrane makes synovial fluid, which has a lubricating function.

T-lymphocyte activation Activation of one type of white blood cell that attacks virus-infected cells, foreign cells, and cancer cells. T cells also produce a number of substances that regulate the immune response.

telomerase A reverse transcriptase that synthesizes telomeric repeat sequences at the ends of chromosomes from its own RNA template.

telomeres or chromosome ends or telomere caps; Repeats of simple-sequence telomeric (TEL) DNA that maintain

the ends of linear chromosomes, counteracting the tendency of a chromosome to be shortened during each round of replication. See *telomerase*. The caps are complexes of protein and DNA.

teratoma A type of germ cell tumour that may contain several different types of tissues, such as hair, muscle, and bone. Teratomas occur most often in the ovaries in women, the testicles in men, and the tailbone in children. Not all teratomas are malignant.

transcription factor A protein that indirectly or directly regulates the activity of RNA polymerase.

transcription The synthesis of an RNA molecule from a DNA template.

transfection Transient or stable introduction of foreign DNA or RNA into eukaryotic cells.

transformation of bacteria The introduction of foreign DNA or RNA into bacterial cells

transforming growth factor □ (TGF-□) A polypeptide growth factor that generally inhibits animal cell proliferation.

translation The synthesis of a polypeptide chain from a mRNA template.

tumour Any abnormal proliferation of cells.

tumour necrosis factor (TNF) A polypeptide growth factor that induces programmed cell death.

tumour suppressor gene A gene whose inactivation leads to tumour development.

tumour virus A virus capable of causing cancer in animals or humans.

ubiquitin A highly conserved protein that acts as a marker to target other cellular proteins for rapid degradation.

vector A DNA molecule used to direct the replication of a cloned DNA fragment in a host cell.

yeast artificial chromosome (YAC) A vector that can replicate as a chromosome in yeast cells and can accommodate very large DNA inserts (hundreds of kb).

zygote A fertilized egg.

Appendix B: Cloning procedures

In this work, four plasmid constructs in total were made by the author, of which two shared a common insert. Three cloning procedures are illustrated below. Standard molecular biology methods were applied, and these have been described in the previous sections.

The pFastBACHTa/COAS2 plasmid was created as outlined in figure A1, with 1) PCR with the COAS2Upper- and LowerBac primers to create a blunt-ended product which was phosphorylated using T4 kinase and ATP. In 2), the FastBACHTa plasmid was dephosphorylated after digestion with the *Ehe*I restriction enzyme, and following T4 ligase-mediated ligation at 14°C over night, the construct was heat-shock transformed into DH5α cells, and positive clones were further screened and treated as outlined in the methods section to yield a purified recombinant plasmid. Dr. D. J. Warren further used this in the Bac-to-Bac system for COAS2 production in insect cells.

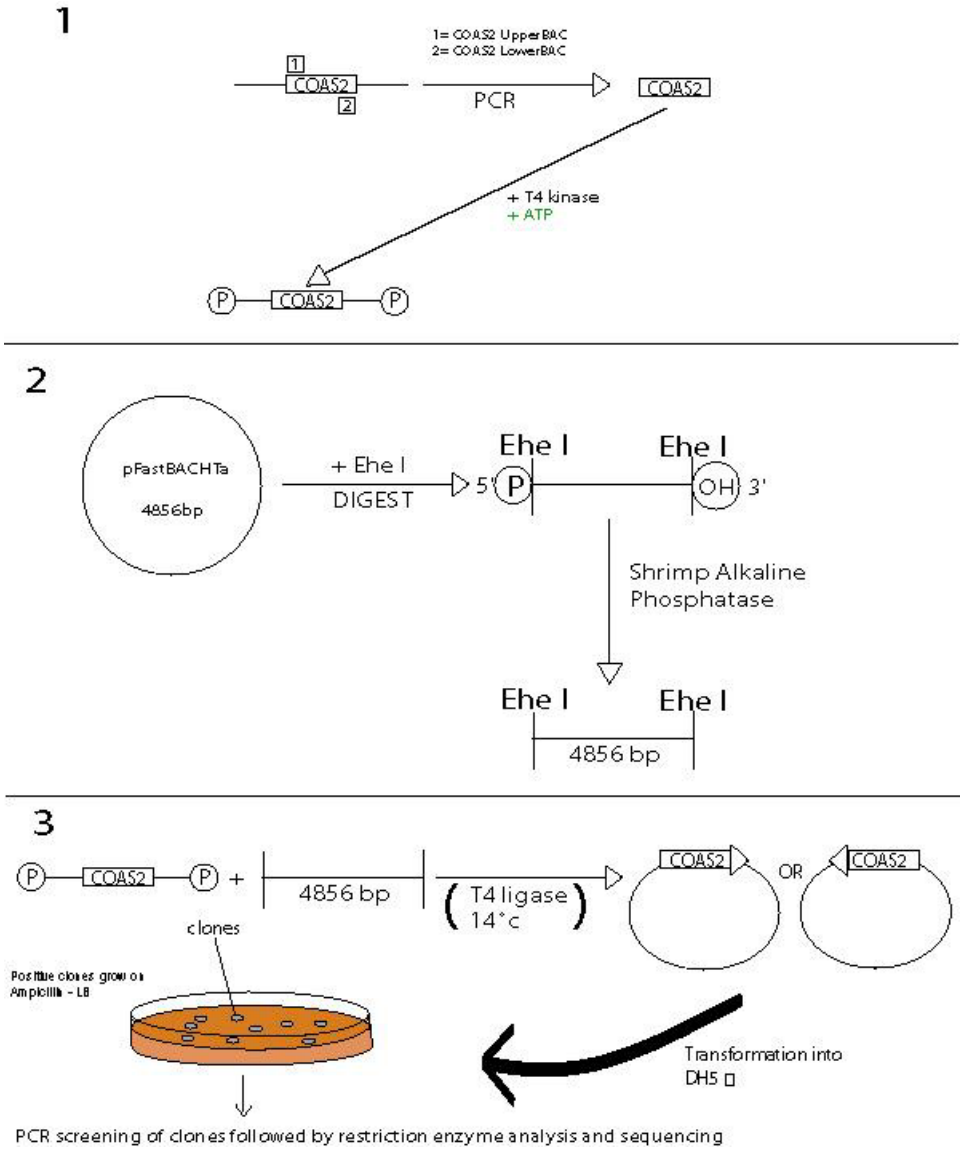


Figure A.1: Illustration of the cloning procedure for the pFastBACHTa/COAS2 plasmid.

The pcDNA5FRT/COAS2MYCHIS$_6$ and pcDNA5FRT/To/COAS2-MycHIS$_6$ plasmids were created as outlined in figures A2-3, with (1) PCR with the SfuILower- and XbaIUpper primers to create a product with overhangs that was gel purified, verified by agarose gel electrophoresis and digested using the *Sfu*I and *Xba*I restriction enzymes (2). The pcDNA3.1 plasmid was also digested with these enzymes, and ligation using T4 ligase over night at 16°C gave a construct that subsequently was heat-shock transformed into JM109 cells (3). Positive clones were further screened and treated as outlined in the methods section to yield a purified recombinant plasmid. The *COAS2MYC-HIS$_6$* was PCR-amplified using the COAS2BamHI and HisTagCOAS2StopXhoI primers, gel purified and verified for proper size using agarose gel electrophoresis (4). Both PCR product and the pcDNA5FRT and pcDNA5FRT/To plasmids were digested using *Bam*HI and *Xho*I restriction enzymes (5), further ligated and

heat-shock transformed into JM109 cells as indicated for the pcDNA3.1COAS2MYCHIS$_6$ B recombinant plasmid (6). Positive clones were further screened and treated as outlined in the methods section to yield two purified recombinant plasmids. These would be used for the creation of stably transfected Flp-In cell lines.

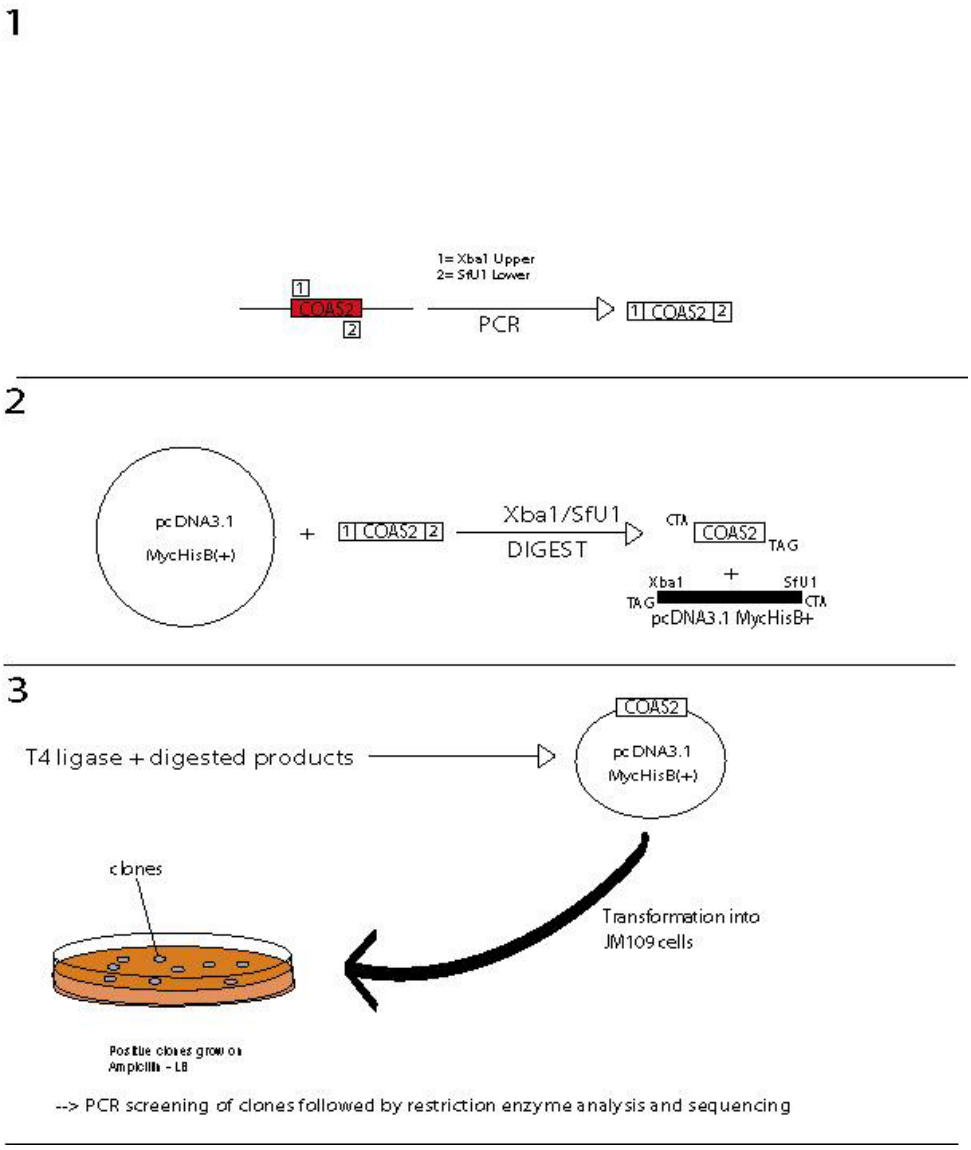


Figure A.2: Illustration of the cloning procedure for the pcDNA3.1MYCHIS$_6$/COAS2 plasmid.

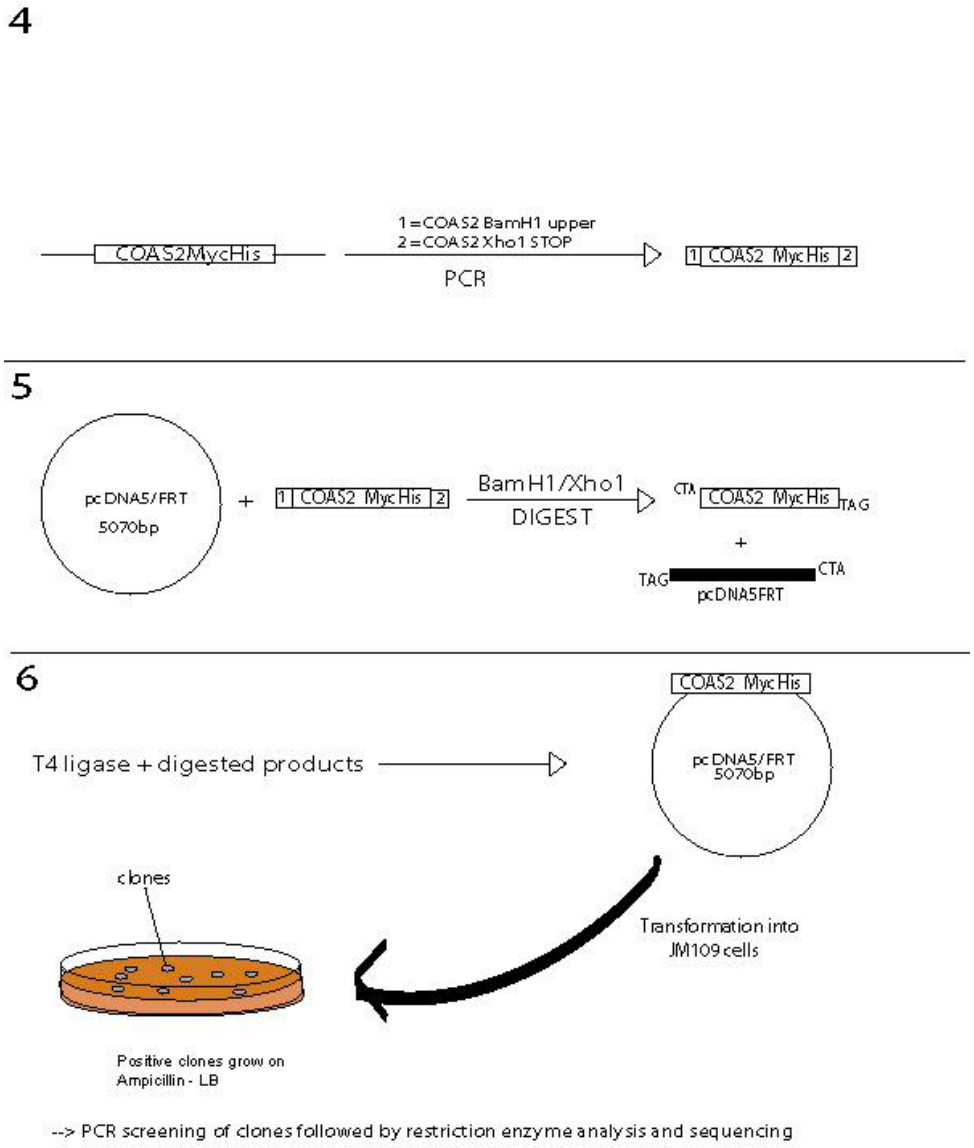


Figure A.3: Illustration of the cloning procedure for the pcDNA5FRT- and FRT/To/COAS2 plasmids.

Appendix C: Materials and equipments

Chemicals

Product	Supplier
[γ-^{32}P]-dATP, 3000 Ci/mmol	Amersham Pharmacia Biotech
[α-^{32}P]-dCTP, 3000 Ci/mmol	Amersham Pharmacia Biotech
Ammonium acetate (NH_4Ac)	Merck
Ammonium persulfate	BIO-RAD, cat. no 161-0700
Ampicillin	Bristol-Meyers Squibb
Agarose	Gibco BRL, Life Technologies
Aprotinin	Sigma, cat. no A4529
Array slidehyb#3	Ambion, 8863
Bacto-agar	Difco Laboratories
β-mercaptoethanol	Sigma
Blasticidin S	Invitrogen, 46-1120, lot no 1188354
Blot Qualified Bovine Serum Albumin	Promega, W384A
Bromophenol blue	Merck
Bovine serum albumin (BSA), 10 x	New England BioLabs
Chloroform	Merck
Coomassie Brilliant Blue	Sigma, B-0630
Cot-1 DNA	Invitrogen
Cy3, NHS-ester leaving group	Amersham Biosciences, PA 23001
Cy5, NHS-ester leaving group	Amersham Biosciences, PA 25001
Diethylpyrocarbonate (DEPC)	Sigma Chemical
dNTPs	Amersham Pharmacia Biotech
dNTPs-C	Amersham Pharmacia Biotech
Dithiotreitol (DTT)	Gibco BRL, Life Technologies
Dimethyl sulphate oxide, DMSO	Sigma, D2650
Dry milk powder (Molico)	Nestlé
Dulbecco's Modified Eagle Medium	BioWhittaker, BE 12-614F
18 S oligo	Amersham Pharmacia Biotech
Ethylene diamine tetra acetate (EDTA)	Merck
Ethidium bromide (EtBr)	Sigma Chemicals
Ethanol, 75%, 95%, abs. EtOH	Arcus
Ficoll 400	Pharmacia, 170400-02
Foetal calf serum, FCS	PAA, A15-043
Formaldehyde, 37%	Merck
Formamide	Merck
FuGene6 Transfection reagent, 0,4 ml	Roche Biosciences, 1815091, lot.no 93051021/93117823
Glucose D+	Merck
Glutamax 1 100x	Invitrogen, 35050-038
Glycine	Merck, IC32046601
Glyserol	Merck
Hydrochloric acid, HCl	Merck
HEPES	Sigma, N9136
Hygromycin B	Invitrogen, lot no A6694
IPTG	Gibco BRL, Life Technologies
Isopropanol	Arcus products
Imidazole	Merck
Kanamycin	Pedersen & Sønn A/S
Leupeptin Hemisulfate,	Sigma, L2884
Magnesium chloride ($MgCl_2$)	Merck
Magnesium dichloride ($MgCl_2$), 25mM	Perkin Elmer

Mowiol	Calbiochem
$NaH_2PO_4 \cdot H_2O$	Merck
NaOH	Merck
Nonidet P-40	USB corporation, 19628
Orange G	Sigma 01625
Penicillin	Invitrogen
Penicillin-Streptomycin for cell culture	BioWhittaker, DE17-602E
Pepstain A	Sigma P4265
PMSF	Sigma, P7626
Poly $d(A)_{40-60}$	Amersham Pharmacia Biotech
Potassium acetate (Kac)	Merck
Potassium chloride (KCl)	Merck
Protein G agarose	Sigma
RNAseZAP	Sigma R-2020
Saponin (extract from the Quillaja tree)	Sigma
Sephadex G50 Medium	Amersham Pharmacia Biotech
Sodium acetate (NaAc)	Serva
Sodium chloride (NaCl)	Merck
Sodium dodecyl sulfate (SDS)	Bio Rad Laboratories
Sucrose	Sigma, S-0389
Tris-(hydroxymethyl)-aminomethane base	Merck, 1083822500
Trizol reagent	Invitrogen, 15596-018
Trypsin/EDTA	Bio-Whittaker
$Tween^{20}$	Merck, S3830084
Vectashield Mounting medium with DAPI	Vector laboratories Inc. H-1200 Lot 1213
Xylene orange	Sigma
Zeocin 100mg/ml	Invitrogen, lot.no 3000038

Instruments

Product		Supplier
Automatic pipettes		Finnpipette
Centrifuges:		
Tabletop: Mikro 20		Hettich Zentrifugen
Biofuge Pico		Heraeus Intruments
Floor: Megafuge 1.0		Heraeus Sepatech
Sorvall RC 5C Plus		Sorvall
Electrophoration:	BTX Pulse Generator	BTX
	Cuvettes Plus 4 mm	BTX
Electroporesis:	Power PAC 300	BioRad
	Gene Pulser II apparature	BioRad
Flow cytometer	FACSCalibur	Becton Dickinson
Homogeniser:	Ultra Turrax T8	IKA labortechnik
Humidified incubator		FormaScientific
Hybond N+ nylon membrane		Amersham Pharmacia Biotech
Hybridization oven		Hybaid
Hypercassette		Amersham Life Science
Incubator (37 ^oC)		Fermaks
Microscope:	Axiovert S100	Zeiss
	LEICA DM RE confocal microscope	LEICA
Microarray scanner		Agilent
Nitrocellulose filter membrane, 0,45 µm		BIO-RAD
PCR machine:	GeneAmp PCR System 9700/2400	PE Applied Biosystems
	Programmable Thermal Controller	MJ Reseach, Inc
pH-meter, PHM 80 Portable		Radometer Copenhagen

Platform Rocker str6		Bibby Stuart
Immobilon-P Polyvinylidene fluoride protein filter, 0,45µm		Millipore, Cat.no IVPH 00010
Sequencing: MegaBACE 500		Amersham Biosciences
Shaker		Gallenkamp Orbital
Spectrophotometric: Gene Quant II		Pharmacia Biotech
	UV1201 UV-VIS spectrophotometer	Bergman
Tubes:	Eppendorf tubes 1.5 ml	Eppendorf
	MicroAmp strips	Perkin Elmer
	GeneAmp PCR tubes 0.5 ml, 0.2 ml	Perkin Elmer
	10 ml polypropylene Falcon tubes	NUNC
	15 ml polypropylene tubes	Sarstedt
	50 ml polypropylene tubes	Sarstedt
	30 ml centrifuge tubes w/lids	Sarstedt
Ultrasonic homogenizer 4710		Cole-Parmer
Ultraviolet Illumination GelDoc 1000		BioRad
UV-crosslinker GS Gene linker		BioRad
Vortex MS2 Minishaker IKA		Tamro lab
Weights:	Sartorius 300,00g	Kebo-Bredal A/S
	SauterAR1014	Nerliens Kemisk-tekniske selskaper
X-OMAT UV film		Kodak

Primers

Name of primer	Sequence, 5'→3'
M13 Forward	CAC ACA GGA AAC AGC TAT G
M13 Reverse	GTT GTA AAA CGA CGG CCA GTG
T7	TAA TAC GAC TCA CTA TAG GG
T3	CAA TTA ACC CTC ACT AAA GG
Sp6	ATT TAG GTG ACA CTA TAG
COAS2BamHI	CGG GAT CCC GAC CAT GGT CAA CTC CGT CGT C
HisTagCOAS2StopXhoI	CCG CTC GAG CGG TTA TCA ATG GTG ATG GTG ATG AT
COAS2XbaIUpper	GCT CTA GAG CAT GGT CAA CTC CGT CGT CTT T
COAS2SfUILower	TTT TTG TTC GAA TCC GAA TTG TCC ACA GTC AGC AAT G
COAS2UpperBAC	GTC AAC TCC GTC GTC TTT TTT G
COAS2LowerBAC	TCA TTA GAA TTG TCC ACA GTC AGC
LICCOAS2Start	GAC GAC GAC AAG ATG GTC AAC TCC GTC GTC
LICCOAS2Stop	GAG GAG AAG CCC GGT TCA TTA GAA TTG TCC ACA GTC AGC AAT

Enzymes and buffers

All enzymes and buffers were from New England Biolabs (NEB) unless otherwise stated. "+BSA" signifies the addition of bovine serum albumin to 0,1 mg/ml.

Enzyme	Buffer	Supplier
Restriction enzymes		
AflIII	3+BSA	
BamHI	Neb buffer for BamHI + BSA	
EcoRI	Neb buffer for EcoRI + BSA	
EcoRV	3+BSA	
HindIII	2	
KpnI	1	
NotI	3	
SacI	1	
SfuI/BstBI	4	

XbaI	2+BSA	
XhoI	2	
Thermostable enzymes		
"DaveTaq", produced at DNR, 5 U/µl	In-house PCR buffer without $MgCl_2$	David J. Warren
"DNRTaq", produced at DNR, 5 U/µl	In-house PCR buffer without $MgCl_2$	David J. Warren
Pfu, produced at DNR, 4 U/µl	In-house Thermopolymerase buffer	David J. Warren
Various enzymes:		
DNA polymerase I (Klenow)		
Platinum®Taq 5 U/µl	PCR Buffer w/o $MgCl_2$	Invitrogen, 10966-026
T4 DNA ligase	NEB buffer for T4 ligase	
T4 Polynucleotide Kinase (PNK)	NEB buffer for T4 PNK	

NEB Buffer composition

BSA dilution buffer (100x): (20 mM, KPO_4, 50 mM NaCl, 0,1 mM EDTA, 5% (v/v) glycerol)
NEB#1: (10 mM Bis Tris Propane HCl, 10 mM $MgCl_2$, 1 mM DTT)
NEB#2: (50 mM NaCL, 10 mM Tris-HCl,10 mM $MgCl_2$, 1 mM DTT)
NEB#3: (100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT)
NEB#4: (50 mM KCH_3CO_2, 20 mM Tris-CH_3CO_2, 10 mM $Mg(CH_3CO_{2)2,}$ 1 mM DTT)
BamHI-buffer: (150 mM NaCL, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT)
EcoRI-buffer: (50 mM NaCL, 100 mM Tris-HCl, 10 mM $MgCl_2$, 0,025% (v/v) Triton-X-100)
T4 PNK buffer: (70 mM Tris-HCl (pH 7,6), 10 mM $MgCl_2$, 5 mM DTT, add ATP to 1 mM)
T4 ligase-buffer: (50 mM Tris-HCl pH 7,5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 µg/ml BSA)

Commercial kits

Kit	Supplier
DYEnamic™ ET dye terminator cycle sequencing kit (MegaBACE)	Amersham Pharmacia Biotech, US81090
Fairplay Micro Array labeling kit, 30 reactions	Stratagene
Jetquick Plasmid Midiprep Spin Kit	Genomed, 210050
Jetquick Plasmid Miniprep Spin Kit 50	Genomed, 200050
Mycoplasma detection kit	ATCC, 90-1001K
QIAquick PCR purification Kit	Qiagen, 28106
QIAquick Gel Extration Kit	Qiagen, 28704
Qiagen EndoFree plasmid maxiprep kit	Qiagen, 12362
ECL kit	Amersham Biosciences, RPN2209
SuperSignalWestPico ECL kit	Pierce, 34078
RTS100 protein expression kit	Roche Biosciences, 3186418
Wizard SV PCR purification kit	Promega, 9281
Microcon YM-30 column	Amicon

Antibodies

Antibody	Supplier	Dilution (range)
Goat anti-chicken filamin, whole serum	Sigma, F2762 Lot. no 3944818	1:40, PBS[11]
Mouse anti-PentaHis, IgG1	Qiagen, 34660 Lot. no 11549390	0,1-2,0 µg/ml in PBS[12]
Secondary antibody rabbit-α-mouse	DAPI, P0161	1:5000 in 5% dry milk (w/v)
Rhodamine/TRITC goat anti-mouse IgG (H+L)	Jackson Immunoresearch, 115-025-146, Lot. no 57054	1:100 in PBS[10]
Anti-His_6	Roche Biosciences, 1922416, lot. no 85272825,	0,1-2,0 µg/ml in TBS[13]
Mouse anti-His_6-HRP	Roche Biosciences, 1965085	As the non-HRP conjugated
Mouse anti-Myc9E10	Mr. David J. Warren	0,1-2,0 µg/ml in TBS[12]
Rhodamine donkey anti-goat #522	Jackson Immunoresearch	1:100 in PBS[10]

Plasmids and markers

Plasmid	Supplier	Selection
pET3a and pET30EkLIC	Novagen	Ampicillin, Kanamycin
PCMVsport6 and PFastBACHTa	Invitrogen	Ampicillin, Ampicillin/Geniticin
pcDNA3.1MycHis B (+)	Invitrogen	Ampicillin/Neomycin
pcDNA5FRT and FRT/To	Invitrogen	Ampicillin/Hygromycin
POG44 and pFRT*lac*Zeo	Invitrogen	Ampicillin
PT7T3DpacI, pOTB7 and pBS KS+	Stratagene	Ampicillin, Chloramphenicol and Chloramphenicol
PEGFP-N1	Clontech	Kanamycin

Molecular weight markers	Supplier
ϕx 174/HaeIII	Invitrogen
λ DNA/HindIII	NEB, digest made in-house
λ DNA/BstEII	NEB, digest made in-house
100 bp ladder	NEB, N3231S
1 kb ladder	NEB, N3232S
2 log ladder	NEB, N3200L

Mammalian cell lines

- 293 Human embryonic kidney (HEK) Flp-In, Flp-In T-rex.

These cells have been adenovirus-transfected. The parental 293 HEK has ATCC number CRL 1573. The Flp-In system has been established in these cells (Invitrogen).

- OHS; Cell line established at the Department of Tumor Biology from primary tumour from femur.

[11] 0,05% saponin (w/v)
[12] 3% BSA (w/v)
[13] with 1-5% dry milk (w/v)

Bacterial strains

- ***E. coli* DH5α** - genotype: F^- ϕ80d*lac*ZΔM15 Δ(*lac*ZYA-argF) U169 *deo*R *rec*A1 *end*A1 hsdR17(r_K^-, m_K^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1

- ***E. coli* JM109** - genotype: *end*A1, *rec*A1, *gyr*A96, *thi,* hsdR17(r_K^-, m_K^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F', *tra*D36, *pro*AB, *lac*I^q ZΔM15]

- ***E. coli BL21(DE3)pLysS*** - genotype: F- *ompT hsd*SB (rB- mB-) *gal dcm* (CamR)

Appendix D: Media and solutions

Media

LB medium (Luria-Bertani Medium)

10 g bacto-tryptone
5 g yeast extract
10 g NaCl
1 l dH_2O
Adjust the pH to 7.0 with 5 M NaOH, sterilise by autoclaving.

For LB-plates add 15 g/l Bacto Agar before autoclaving. Antibiotics are added after autoclaving when the medium holds 60 ^oC or less. Ampicillin: 50 mg/ml, Kanamycin: 10 mg/ml, Chloramphenicol: 50 mg/ml.

SOC-medium

20 g bacto-trypton
5.0 g yeast extract
0.5 g NaCl
800 ml dH_2O
1 ml 2.5 M KCl
Adjust the pH to 7.0 with 5 M NaOH, sterilize by autoclaving. When cool, add 5 ml 2 M $MgCl_2$ and 20 ml 1 M solution of glucose

Dulbecco's Modified Eagle Medium (DMEM) for the Flp-In 293 cell line

DMEM
8% Fetal Calf Serum
2mM PenStrep
2 x Glutamax
100 μg/ml Zeocin

DMEM for the Flp-In 293 T-Rex cell line

DMEM with additives as devised for the Flp-In 293 cell line
15 μg/ml Blasticidin

DMEM for Flp-In 293 transfected with gene of interest (positive transfectants)

DMEM with additives as devised for the parental cell lines without Zeocin, nor Blasticidin
0,15 μg/ml Hygromycin B

DMEM for Flp-In 293-T-Rex transfected with gene of interest (positive transfectants)

DMEM as devised for the transfected cell lines
15 μg/ml Blasticidin

Solutions

Agarose gel, 1 %

50 ml 1 x TAE buffer
0.5 g agarose
Boil to dissolve the agarose
Cool to 60ºC, then add 1-2 µl ethidiumbromide (10 mg/ml), swirl to mix and pour into casting tray.

Amidoblack staining solution for protein

1 g Naphtol Blue Black
450 ml MeOH
100 ml acetic acid
ddH_2O to 1 litre

Amidoblack wash solution

900 ml MeOH
20 ml acetic acid
80 ml ddH_2O

Ammonium chloride, NH_4Cl in PBS, 50 mM

133 mg NH_4Cl
PBS to 100 ml
Filter-sterilize the solution by filtration through a 0.2 µm filter.

Bjerrum-Schäfer-Nilsen 1 x blotting-buffer for Western

5,8 g Tris-amino hydroxymethane
2,9 g glycine
ddH_2O to 1 litre

Blotting-buffer "A" for Western with Methanol

25 mM Tris pH8.3
192 mM glycine
20% (v/v) MeOH
ddH_2O
Add MeOH as final ingredient.

50 mM $CaCl_2$:Glyserol (85:15)

85 ml 50 mM $CaCl_2$
15 ml 99 % glycerol

Church hybridization solution (RNA)

500 mM Sodium phosphate pH 7.2, 7 % SDS, 1 mM EDTA
250 ml 1 M Sodium phosphate pH 7.2
175 ml 20 % SDS
1 ml 0.5 M EDTA
DEPC-dH_2O to 500 ml

Church wash solution (RNA)

40 mM Sodium phosphate pH 7.2, 1 % SDS
40 ml 1 M Sodium phosphate pH 7.2
50 ml 20 % SDS
DEPC-dH_2O to 1000 ml
For DNA: Church hybridization solutions as for RNA except ddH_2O.

Coomassie Brilliant Blue staining solution for protein

1 g Coomassie Brilliant Blue
500 ml MeOH
100 ml glacial acetic acid
ddH_2O to 1 litre

Coomassie Brilliant Blue fix/wash solution

400 ml MeOH
100 ml glacial acetic acid
ddH_2O to 1 litre

DEPC treated water

0,1% DEPC
ddH_2O
Let the solution stand for 1 hour after DEPC-addition, then autoclave. Shake container while still hot for final inactivation of excess reagent. It is important to be sure that DEPC has been inactivated as it will modify RNA and inactivate enzymes such as reverse transcriptases and kinases.

Dithiothreitol, DTT, 1 M

3.09 g dithiothreitol
20 ml 10 mM sodium acetate pH 5.2
Sterilize by filtration, aliquot and store at –20 ^oC

DNAse I-buffer

10 mM Tris-Cl pH 7.5
150 mM NaCl
1 mM $MgCl_2$
DNAse I
DNAse I may be diluted in this buffer in a concentration range from 50 to 500 ng/ml.

EDTA, 0.5 M pH 8.0

186.1 g disodium EDTA·$2H_2O$
dH_2O
Dissolve EDTA in 800 ml water, then stir vigorously on a magnetic stirrer. Adjust the pH to 8.0 with NaOH, finally add dH_2O to 1000 ml and autoclave

Etidiumbromide, EtBr, 10 mg/ml

1 g ethidium bromide
dH_2O to 100 ml
Stir on a magnetic stirrer until the dye has dissolved. Aliquot. Store at 4 °C, dark.

Gel loading dye #1, 6x, for DNA

0.25 % (w/v) bromphenol blue
0.25 % (w/v) xylene cyanol FF
30 % (v/v) glycerol in water

Gel loading dye #2, 6x FEBX, for DNA

15% Ficoll 400
0,25 % Orange G Sigma cat 01625

Gel loading dye, 1,3x, for RNA

20 mM NaH_2PO_4 pH 6.6
66 % formamide
8,6 % formaldehyde
0,07 % xylene-orange
0,7 µg/ml EtBr
ddH_2O
Mix in a ventilated area, preferably under a safety hood. Aliquot and store at -70°C.

Gel loading buffer, 2x, for protein

2 ml sH_2O
1 ml 1 M Tris-Cl pH 6.8
4 ml 50 % glycerol
1 ml 2 % Brom-phenol blue (BFB)
2ml 20% SDS (w/v)

Heat SDS at 37ºC for 15 minutes in Tris and water to dissolve. Mix, aliquot and store at -20ºC.

Gel loading buffer, 6x, for protein

15 % SDS,
50 % glycerol
0,3 M Tris-Cl pH6.8
25 % beta-mercaptoethanol
Brom-phenol blue (BFB)
Heat SDS at 37ºC for 15 minutes in Tris and water to dissolve. Mix, aliquot and store at -20ºC.

Isopropyl-beta-D-thiogalactopyranoside, (IPTG), 100 mM

238 mg IPTG
10 ml dH_2O
Sterilize by filtering through a 0,45 µm filter.

Lysisbuffer for purification of E.coli under native conditions

As described in the QiaExpressionist-protocol collection:

50 mM NaH_2PO_4
300mM NaCl
10mM imidazole
Adjust to pH 8.0 with NaOH.

Lysozyme, 10 mg/ml

Dissolve to 10 mg/ml in 10mM Tris-Cl pH 8.0 immediately before use.
Verify pH≥8.0 to ensure proper lysozyme activity in solution.

Magnesium chloride $MgCl_2$, 1 M

203 g $MgCl_2 \cdot 6H_2O$
dH_2O to 1000 ml

NT-stop solution for gel filtration

20 mg/ml blue dextran, 1 % SDS, 40 mM EDTA
600 µl 25 mg/ml blue dextran
50 µl 10 % SDS
50 µl 0.5 M EDTA

OLB°-buffer, 5 x

250 µl 1 M Tris-HCl pH 7.6
25 µl 1 M $MgCl_2$
500 ml 2 M HEPES pH 6.6
3.5 µl β-mercaptoethanol
220 µl dH_2O

Paraformaldehyde, 4 %

4 g paraformaldehyde
10 ml dH_2O
Heat to 60°C, add 2-6 droplets 1 M NaOH until the solution becomes clear, then cool. Sterilize by filtration. Add PBS to 100 ml, adjust pH to 7.4. Store at 4°C, dark for 14 days maximum.

Phosphate-buffer saline (PBS), 1 x

2.7 mM KCl, 1.4 mM KH_2PO_4, 137 mM NaCl, 4.3 mM Na_2PO_4
200 mg KCl
200 mg KH_2PO_4
8.0 g NaCl
1.15 g $Na_2PO_4 \cdot 7H_2O$
800 ml dH_2O
Adjust pH to 7.2 with 1 M NaOH
dH_2O to 1000 ml
Sterilize by filtration through a 0,2-0,45 µm filter.

Potassium acetate (KAc), 5 M

491 g KAc
dH_2O to 1000 ml, autoclave.

RNAse A, 5 mg/ml

50 mg RNase A
10 ml 10 mM NaAc pH 5.2
Incubate 5 minutes at 100 ^oC, then cool to RT.
Adjust pH to 7.4 with 1 M HCl, aliquote and store at -20 degrees.

Sodium Acetate, NaAc, 3 M pH 5.2

408 g $NaAc{\cdot}3H_2O$
800 ml dH_2O
Adjust pH with glacial acetic acid
dH_2O to 1000 ml, autoclave.

Sodium hydroxide, NaOH, 10 M

400 g NaOH
dH_2O to 1000 ml, dissolve gently due to heat generation, preferentially in glass equipment. Filter-sterilise.

Sodium phosphate, 1 M pH 7.2

134 g $Na_2HPO_4{\cdot}7H_2O$
800 ml dH_2O
Adjust pH with 85 % H_3PO_4
dH_2O to 1000 ml, autoclave.

Sodium phosphate, 0.2 mM pH 6.6

27.6 g $NaH_2PO_4{\cdot}H_2O$
800 ml DEPC-H_2O
Adjust pH with 10 M NaOH, add DEPC-H_2O to 1000 ml and finally autoclave.

Standard saline citrate, SSC, 20 x

3 M NaCl, 0.3 M sodium citrate
175 g NaCl
88.2 g $Na_3citrate{\cdot}2H_2O$
800 ml DEPC-dH_2O
Adjust pH to 7.0 with 1 M HCl
dH_2O to 1000 ml, autoclave.

TAE-buffer, 50 x

40 mM Tris-acetate, 1 mM EDTA
242 g Tris (hydroxymethyl aminomethane) base
57.1 ml glacial acetic acid
100 ml 0.5 M EDTA pH 8.0
dH_2O to 1000 ml
Mix water, Tris and acetic acid in a safety hood. Finally add EDTA to adjust pH, then autoclave.

TBE-buffer, 10 x

45 mM Tris-borate, 1 mM EDTA
108 g Tris (hydroxymethyl aminomethane) base
55 g boric acid
40 ml 0.5 M EDTA pH 8.0
Add EDTA as final ingredient to adjust pH, autoclave.

TE-buffer 1x

10 mM 1 M Tris-HCl pH 8.0, 1 mM EDTA
10 ml 1 M Tris-HCl pH 8.0
2 ml 0.5 M EDTA
dH_2O to 1000 ml, autoclave.

Tris-buffered saline, TBS for western blotting

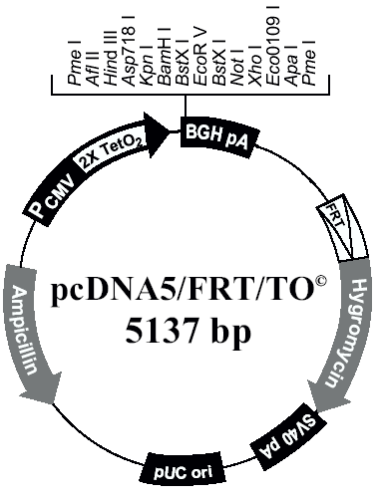
10 mM Tris-Cl, pH 7.5
150 mM NaCl

Western running buffer 1 x

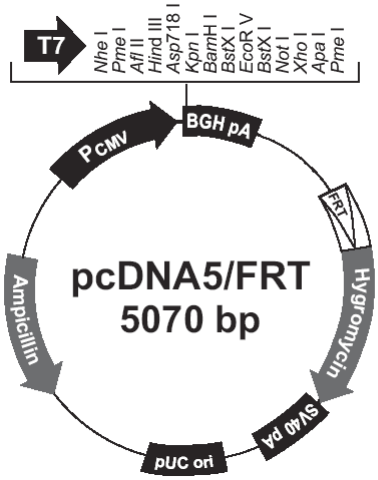
3,02 g Tris (hydroxymethyl aminomethane) base
14,4 g glycine
5 ml 20% SDS

Appendix E: Plasmid maps

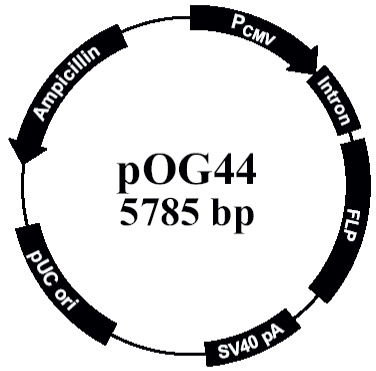
pcDNA5/FRT/TO



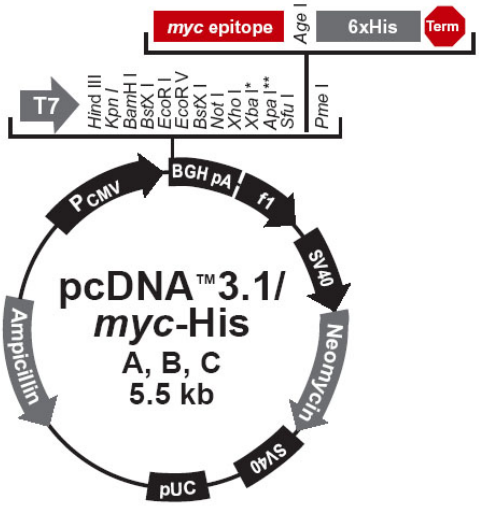
pcDNA5/FRT



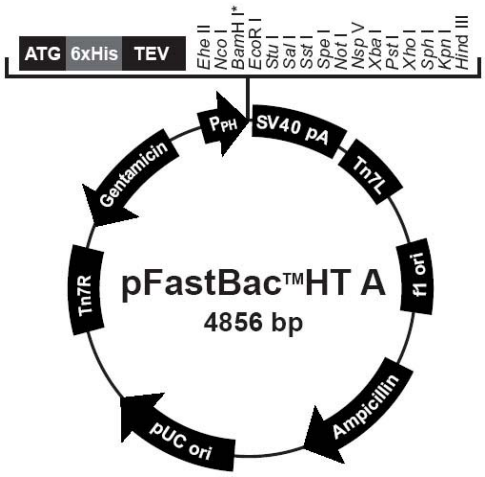
pOG44



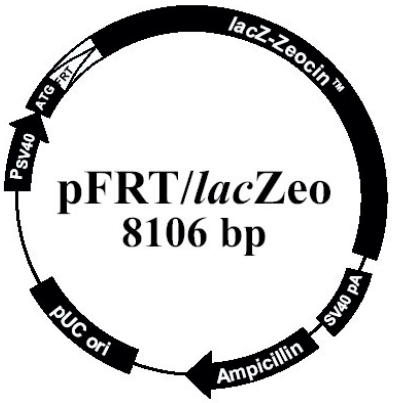
pcDNA3.1/*myc*-His A,B,C



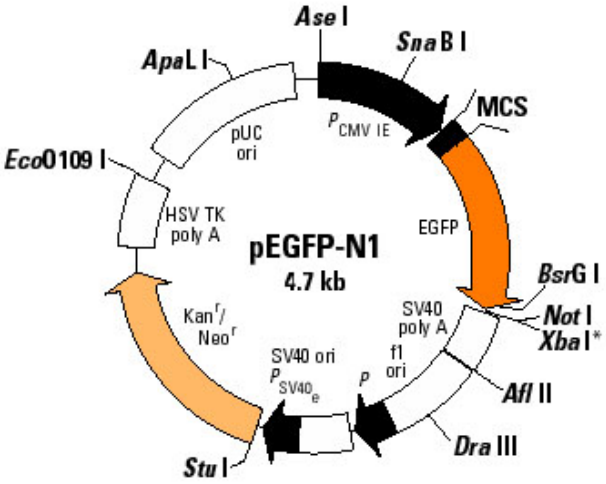
pFastBacHT A



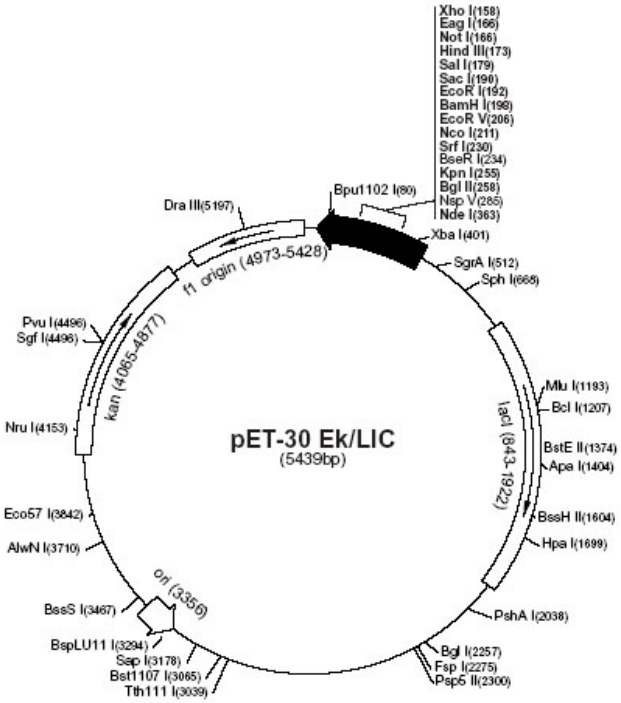
pFRT/*lac*Zeo



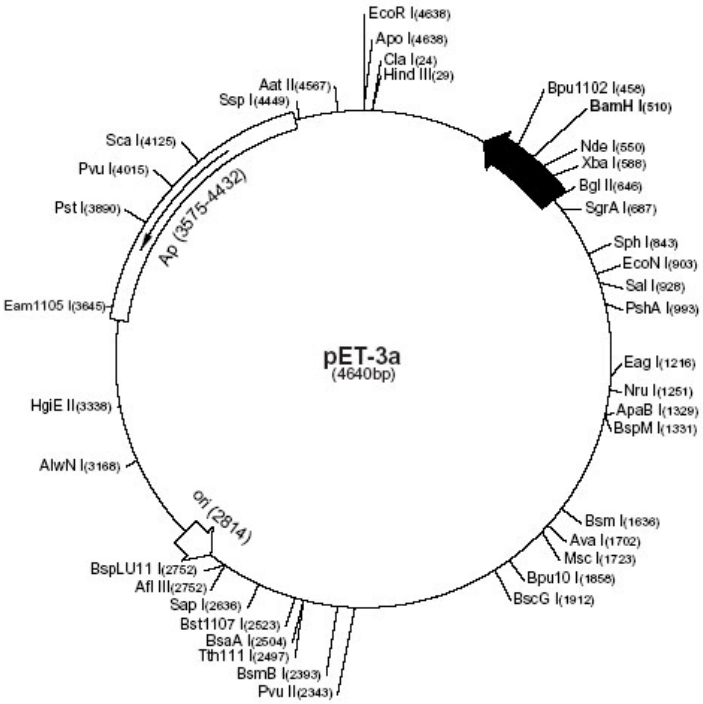
pEGFP-N1



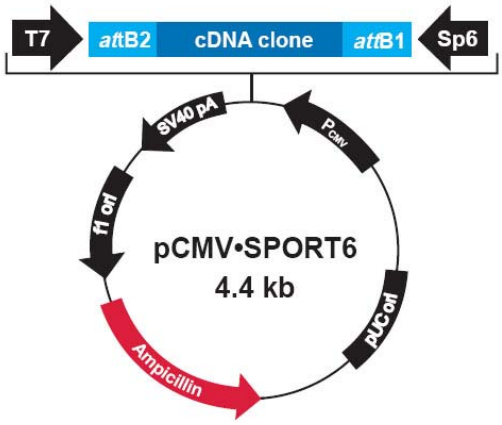
pET-30 Ek/LIC



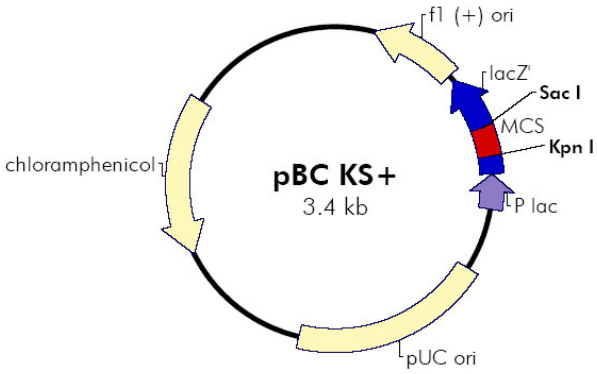
pET-3a



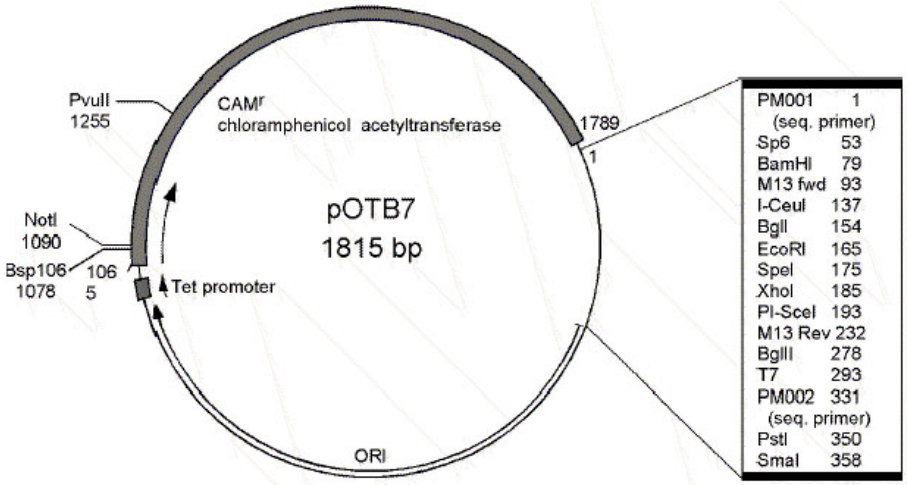
pCMV-SPORT6



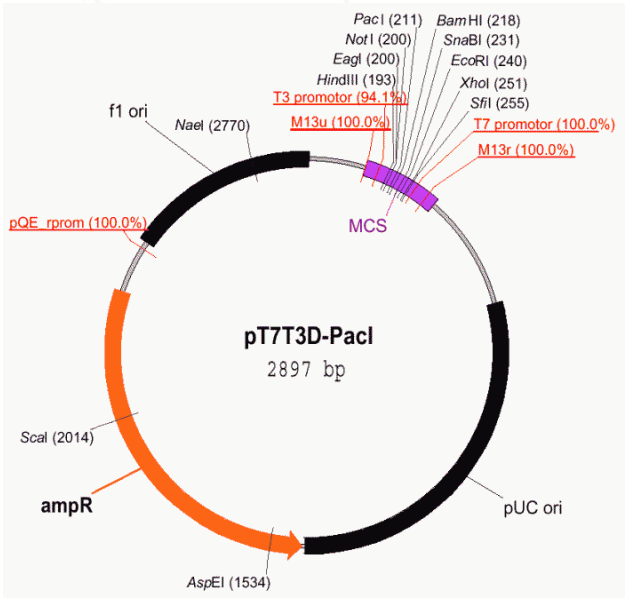
pBluescript KS+



pOTB7



pT7T3D-PacI



APPENDIX F: Internet sites

In addition to URLs given here, many links to interesting websites can be found at http://folk.uio.no/leonardm, http://test.radium.no/myklebost, http://folk.uio.no/mskarn, and http://research.radium.no/institute

http://www.atcc.org	The global bioresource center.
http://cgap.nci.nih.gov	The cancer genome anatomy project.
http://www.nih.gov/news/stemcell/scireport	Stem Cell info from NIH, USA.
http://stemcells.nih.gov/infoCenter/stemCellBasics.asp	An introduction to stem cells.
http://www.molmine.com	Software solutions for microarray data analysis.
http://alba.uio.no/base	The BASE microarray software environment
http://www.bioinfo.no	The FUGE bioinformatics platform homepage.
http://www.enzyme-halle.mpg.de	Protein folding and PPI-material.
http://www.biocarta.com	Models of eukaryotic signal pathways available online.
http://us.expasy.org	The **Ex**pert **P**rotein **A**nalysis **Sy**stem proteomics server.
http://www.rcsb.org/pdb	The Protein Data Bank
http://www.pubcrawler.ie	Updates the contents of Medline and GenBank.

Database accession numbers

Database	Accession number
NCBI Nucleotide and protein ID	COAS2: GI:4826471; NP_839944 L-SOX5: GI:23308714 SOX6: GI:13435017 SOX9: GI:4557852
REFSEQ-accession number	COAS2: NM_178230 L-SOX5: NM_152989
(ExPaSy) pfam (protein family)	COAS2: 160
LocusLink	COAS2: 164022 COL2A1:1280
UniGene	COAS2: Hs.523569